(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) Publication number: **0 426 195 A1**

(12) EUROPEAN PATENT APPLICATION

(21) Application number: 90121052.6

(22) Date of filing: 02.11.90

(51) Int. Cl.5: **C12N 15/40**, C12N 15/82, C12N 15/11, A01H 5/00, C12Q 1/68

(30) Priority: 03.11.89 US 431259
05.12.89 US 446024

(43) Date of publication of application:
08.05.91 Bulletin 91/19

(84) Designated Contracting States:
AT BE CH DE DK ES FR GB GR IT LI LU NL SE

(71) Applicant: **ZAADUNIE B.V.**
Westeinde 62
NL-1601 BK Enkhuizen(NL)

(72) Inventor: **Goldbach, Robert Willem**
Laarweg 9
NL-6721 DA Bennekom(NL)
Inventor: **Peters, Dirk**
Edelmanlaan 4
NL-6703 Wageningen(NL)
Inventor: **Gielen, Johannes Jacobus Ludgerus**
Jan Gooskaai 73
NL-1602 GC Enkhuizen(NL)
Inventor: **De Haan, Petrus Theodorus**
Prins Bernhardstraat 32
NL-6668 AH Randwijk(NL)
Inventor: **Kool, Arnoldus Johannes**
Kerkeland 74
NL-1602 LJ Enkhuizen(NL)
Inventor: **Van Grinsven, Martinus Quirinus Joseph Marie**
Wezenland 5
NL-1602 MA Enkhuizen(NL)

(74) Representative: **Smolders, Walter et al**
c/o Sandoz Technology Ltd., Lichtstrasse 35
CH-4002 Basel(CH)

(54) **Improvements in or relating to organic compounds.**

(57) Recombinant DNA constructs comprising a DNA coding for transcription into an RNA sequence of tospoviruses or into RNA sequences related thereto, the use of such DNA constructs to transform plants having reduced susceptibility to tospovirus infection and probes for the isolation of tospovirus or diagnosis of plant tospovirus diseases

EP 0 426 195 A1

IMPROVEMENTS IN OR RELATING TO ORGANIC COMPOUNDS

Scope of the invention

The present invention relates to plants having reduced susceptibility to infection with tospoviruses, genetic material used to generate this tolerance, probes for the isolation or diagnosis and processes for obtaining such plants and genetic material and probes

Introduction

In general, virus infections have a variety of detrimental effects on e g growth, morphology and yield of plants Also, virus infections often result in higher susceptibility of infected plants to other plant pathogens and plant pests Transmission of plant viruses occurs generally by insect or fungal vectors or mechanically

Plant breeders are trying continuously to develop varieties of crop plant species tolerant to or resistant to specific virus strains Traditionally virus resistance genes are transferred from wild relatives into the commercial varieties by breeding The transfer of an existing resistance from the gene pool of wild relatives to a cultivar is a tedious process in which the resistance gene(s) first have to be identified in a source (donor) plant species and the resistance genes have to be combined with the perfect gene pool of a commercial variety Resistances or tolerances generated in this way are in many cases only active against one or a few strains of the virus in question Also breeding cultivars for resistance to a particular virus species is often limited by a lack of genetic sources for this resistance within the crop species Other approaches to prevent or decrease the effect of virus disease on plants are the use of chemicals or other methods which act against the virus vectors such as e g the use of insecticides, fungicides or good phytosanitary working conditions However, the use of chemicals to combat virus disease by killing the vector is subject to increasing stricter regulations for use imposed by governments, confronting the growers with a decreasing scala of allowed chemical plant-protectants

Alternatively, a system called "cross-protection" can be employed Cross-protection is a phenomenon in which infection of a plant with one strain of a virus protects the plant against superinfection with a second related virus strain Thus, this latter method preferentially involves the use of avirulent virus strains to infect plants, to inhibit a secondary infection with a virulent strain of the same virus However, the use of this natural cross-protection system has several disadvantages The method is very laborious because it requires inoculation of every planted crop, and carries the risk that due to a mutation the former avirulent strain becomes a more virulent strain, thereby creating a disease by itself It is also possible that an avirulent virus strain on one plant species acts as a virulent strain on another plant species

Several studies indicated that the viral coat protein of the protecting virus plays an important role in the cross-protection and that protection occurs when the resident virus and the challenging virus have the same or a closely related coat protein structures

With the recent development of genetic manipulation and plant transformation systems new methods to create virus resistance have emerged Genetically engineered cross-protection is a form of virus resistance which phenotypically resembles the natural cross-protection, but is achieved through expression of the genetic information of the viral coat protein from the genome of a genetically manipulated plant The first successful experiments generating virus resistance by genetic engineering were performed by Beachy *et al* (1985) and Abel *et al* (1986) They showed that expression of the tobacco mosaic virus strain U1 (TMV-U1) coat protein gene from the genome of a transgenic plant resulted in a delay of symptom development after infection with any TMV strain Similar results with respect to coat protein-mediated protection have been obtained for alfalfa mosaic virus (AMV), potato virus X (PVX) and cucumber mosaic virus (CMV)

Although TMV, CMV, AMV and PVX belong to different virus groups, they share a common architecture: in all these virions the viral RNA is a positive strand RNA encapsidated by a viral coat consisting of many individual but identical viral coat proteins

Tospoviruses are essentially different from these plant virions The genus of tospoviruses belongs to the Bunyaviridae. All tospoviruses are thrips transmitted The virus particles are spherically shaped (80-120 nm in diameter) and contain internal nucleocapsids surrounded by a lipid envelope studded with glycoprotein surface projections The multipartite genome consists of linear single stranded RNA molecules of negative or ambisense polarity The terminal nucleotides of these RNA molecules are characterised by a consensus sequence as follows: 5′ AGAGCAAUX ..... GAUUGCUCU 3′, wherein X is C or U Typical members of the tospoviruses are tomato spotted wilt virus (TSWV) and Impatiens necrotic spot virus, also known as

TSWV eyetype

The TSWV virion contains 4 distinct structural proteins: an internal nucleocapsid protein N of 29 kd and two membrane glycoproteins: G_1 of approximately 78 kd and G_2 of approximately 58 kd. In addition, minor amounts of a large protein L of approximately 260 kd have been detected in virus particles. The genome of TSWV consists of three linear single stranded RNA molecules of ±2900 nt (S RNA), ±5000 nt (M RNA) and ±7500 nt (L RNA), each tightly associated with nucleocapsid proteins and a few copies of the L protein to form circular nucleocapsids. A schematic structure outlining most properties of the TSWV virion is given in figure 1. Based on these and other properties TSWV has been classified as a representative of the tospoviruses group. Moreover, TSWV is considered as the typemember of the tospoviruses. Only circumstantial evidence was presented that suggested that a M RNA encoded gene is directly or indirectly involved in the synthesis of the G_1 membrane glycoprotein (Verkleij and Peters, 1983). The coding properties of the other RNA molecules and the polarity of the genomic RNA were still unknown prior to this invention.

As stated earlier, tospoviruses such as TSWV are transmitted by certain thrips species. The vectors of TSWV belong to the family of *Tripidae* and include tobacco thrips (*Frankliniella fusca* (Hinds)), western flower thrips (*F. occidentalis* (Pergande)), common blossom thrips (*F. Schultzei* (Trybom)), chilli thrips (*Scirtothrips dorsalis* (Hood)), *Thrips setosus* (Moulton), onion thrips (*T. tabaci* (Lindeman)) *F. intonsa and T. palmi*. Virus is acquired by the thrips only during their larval stages. Larvae can transmit the virus before they pupate but adults more commonly transmit the virus. Adults can remain infective throughout their life span.

The current distribution of TSWV covering all the continents makes it one of the most widely distributed plant viruses. The virus is widespread in temperate, subtropical and tropical climate zones throughout the world. At least 370 plant species representing 50 plant families, both mono- and dicotyledons, are naturally infected. The TSWV seriously affects the production of food and ornamental crops. Infections of plants with TSWV strains result generally in e.g. stunting, ringspots, dark purple-brown sunken spots, stem browning, flower breaking, necrotic and pigmental lesions and patterns, yellows and non-necrotic mottle, mosaic in greens or even total plant death. Most hosts only exhibit a part of these symptoms. The wide range of symptoms produced by TSWV has complicated the diagnosis and led to individual diseases being given several different names. Also, TSWV symptoms within the same plant species may vary depending on the age of the plant, time of infection during the life-cycle of the plant, nutritional levels and environmental conditions, especially temperature.

Although TSWV has been known for many years, is widely distributed, and causes economically important diseases in crops and ornamentals, limited progress has been made to identify sources for TSWV resistance genes. A multigenic TSWV tolerance has been identified in *Lycopersicon peruvianum*, but this resistance has not been transferred yet to cultivated tomatoes nor has a resistance source been identified for other crop species. The use of natural cross-protection systems to decrease the damage by severe TSWV strains is not well documented. Limited positive results have been reported for tomato and lettuce.

Therefore, the introduction of genetic information conferring resistance to tospovirus infection into plant gene pools by means of genetic manipulation provides breeder and grower a new method to combat tospoviral diseases.

Detailed Description

The present invention provides recombinant DNA constructs comprising a DNA sequence coding for transcription into

a) an RNA sequence of tospoviruses or an RNA sequence homologous thereto

b) an RNA sequence according to a) encoding for a tospovirus protein in which one or more codons have been replaced by their synonyms (i.e. codons corresponding to the same amino acid or termination signal), or a part thereof; or an RNA sequence homologous thereto, or

c) an RNA sequence complementary to an RNA sequence according to a) or b), which DNA is under expression control of a promoter functioning in plants and has a terminator functional in plants.

The DNA sequences defined under a), b) and c) hereinabove, are, hereinafter, for convenience referred to as a "TSWV Related DNA Sequences". A TSWV Related DNA Sequence according to the invention may be modified, if desired, to create mutants or modified sequences homologous to a TSWV Related DNA Sequence from which they are derived using methods known to those skilled in the art such as site-directed mutagenesis. Such mutants or modified coding sequences are therefore within the spirit and the scope of this invention.

The term RNA sequence of a tospovirus refers to a sequence of the S, M or L RNA strand in particular of the S or L RNA strand, preferable of the S RNA strand of a tospovirus

The term RNA sequence homologous to an RNA sequence of a tospovirus wherein a number of nucleotides have been deleted and/or added but is still capable of hybridization to a nucleotide sequence complementary to an RNA sequence of a tospovirus under appropriate hybridization conditions. For the purpose of the invention appropriate hybridization conditions conveniently include an incubation for 16 hours at 42°C, in a buffer system comprising 5 x standard saline citrate (SSC), 0.5% sodium dodecylsulphate (SDS), 5 x Denhardt's solution, 50% formamide and 100 µg/ml carrier DNA (hereinafter the buffer system), followed by washing 3 times with a buffer comprising 1 x SSC and 0.1% SDS at 65°C for approximately one hour each time.

Preferred hybridization conditions for the purpose of the invention involve incubation in the buffer system for 16 hours at 49°C and washing 3 times with a buffer comprising 0.1 x SSC and 0.1% SDS at 55°C for approximately one hour each time. Most preferred hybridization conditions for the purpose of the invention involve incubation in the buffer system for 16 hours at 55°C and washing 3 times with a buffer comprising 0.1 x SSC and 0.1% SDS at 65°C for approximately one hour each time.

The length of the TSWV Related DNA Sequence will i a depend on the particular strategy to be followed, as will become apparent from the description hereinafter. In general, it will be desirable that the TSWV Related DNA Sequence comprises at least 20, suitably 50 or more nucleotides.

The term promoter as used herein refers to the nucleotide sequence upstream from the transcriptional start site and containing all the regulatory regions required for transcription including the region coding for the leader sequence of mRNA (which leader sequence comprises the ribosomal binding site and initiates translation at the AUG start codon).

Examples of promoters suitable for use in DNA constructs of the invention include viral, fungal, bacterial, animal and plant derived promoters functioning in plant cells. The promoter may express the DNA constitutively or differentially. Suitable examples of promoters differentially regulating DNA expression are promoters inducible by disease vectors, such as thrips, e.g. so-called wound inducible promoters. It will be appreciated that the promoter employed should give rise to the expression of a TSWV Related DNA Sequence at a rate sufficient to produce the amount of RNA necessary to decrease the tospovirus susceptibility of the transformed plant. The necessary amount of RNA to be transcribed may vary with the type of plant involved. Particularly preferred promoters include the cauliflower mosaic virus 35S (CaMV 35S) promoter, derivatives thereof, and a promoter inducible after wounding by a disease vector such as thrips, e.g. a wound inducible promoter.

The term terminator as used herein refers to a DNA sequence at the end of a transcriptional unit that signals termination of transcription. Terminators are DNA 3′-non-translated sequences that contain a polyadenylation signal, that causes the addition of polyadenylate sequences to the 3′-end of the primary transcript. Terminators active in plant cells are known and described in the literature. They may be isolated from for example bacteria, fungi, viruses, animals and plants. Examples of terminators particularly suitable for use in the DNA constructs of the invention include the nopaline synthase terminator of *A. tumefaciens*, the 35S terminator of CaMV and the zein terminator from *Zea mays*.

According to the terminology employed in the present specification, an RNA sequence is complementary to another RNA sequence if it is able to form a hydrogen-bonded complex with it, according to rules of base pairing under appropriate hybridization conditions (as defined hereinabove).

The invention also provides a vector capable of introducing the DNA construct of the invention into plants and methods of producing such vectors.

The term vector employed herein refers to a vehicle by means of which DNA fragments can be incorporated in a host organism.

The term plants is used herein in a wide sense and refers to differentiated plants as well as undifferentiated plant material such as protoplasts, plant cells, seeds, plantlets etc. that under appropriate conditions can develop into mature plants, the progeny thereof and parts thereof such as cuttings and fruits of such plants.

The invention further provides plants comprising in their genome a DNA construct of the invention, and methods of producing such plants.

The plants according to the invention have reduced susceptibility to diseases induced by tospoviruses and do not have the disadvantages and limitations of plants obtained by the classical methods as discussed hereinabove.

Examples of plants susceptible to tospoviruses such as TSWV include but are not limited to Ageratum, alfalfa, Amaranthus, Anthirrhinum, Aquilegia, aubergine, beet, Begonia, broad bean, broccoli, brussels sprouts, cabbage, cauliflower, celery, chicory, Chrysanthemum, Cineraria, clover, Cosmos, cowpea cu-

cumber, cyclamen, Dahlia, Datura, Delphinium, endive, Gerbera, Gladiolus, Gloxinia, gourd, groundnut, Hippeastrum, Impatiens, lettuce, melon, Mesembryanthemum, onion, papaya, pea, peanut, pepper, petunia, pineapple, potato, Primula, Saint Paulia, safflower, Salpiglossis, snap bean, soybean, spinach, squash, sugarbeet, sunflower, Tagetes, tobacco, tomato, Verbena, Vinca, watermelon, Zinnia. The invention relates in particular to these listed plants comprising in their plant genome a DNA construct of the invention.

Since TSWV is the typemember of the tospoviruses, the particular features of tospoviruses are hereinafter illustrated employing TSWV as an example.

The S, M and L RNA are single stranded RNA molecules. The S RNA is approximately 2,900 nucleotides long and comprises two genes, one encoding a non-structural protein (NSs) in viral sense, the other one encoding the nucleocapsid protein (N) in viral complementary sense. The intergenic region between the NSs- and N-gene can be folded into a hairpin structure. The 5′-and 3′-terminal sequences of the S RNA are capable of hybridizing to each other such that the first nucleotide is opposite (and complementary) to the last nucleotide of said S RNA strand. We designate hereinafter the double-stranded structure obtained by hybridizing both RNA termini for convenience "pan-handle".

The M RNA strand has approximately 5000 nucleotides. It contains one long open reading frame in viral complementary sense. This open reading frame is translated on polysomes located on the endoplasmic reticulum where the nascent polypeptide chain is cleaved co-translationally to form the spike proteins G_1 and G_2 respectively. Similar to the S RNA the termini of the M RNA strand are complementary to each other and may likewise hybridize to form a "pan-handle".

The L RNA strand has approximately 7500 nucleotides and contains one open reading frame in the viral complementary sense. This open reading frame most probably corresponds with the gene encoding the viral transcriptase with an estimated molecular weight of approximately 260 kd. In some mutant strains shortened L RNA molecules have been found in addition to the wild type, full-length L RNA. These shortened L RNAs however, do always possess the characteristic terminal nucleotide sequences and thus are capable of forming "pan-handle" structures. They are also encapsidated with nucleocapsid protein and included in virus particles. Their presence supresses symptom development resulting in less severe detrimental effects. Hence, these shortened L RNA molecules can be regarded as defective interfering (DI) RNAs, a term known to those skilled in the art.

A preferred embodiment of the invention relates to DNA constructs of the invention coding for transcription into tospovirus-RNA sequences of a "pan-handle", or into tospovirus-RNA sequences homologous thereto.

Another preferred embodiment of the invention relates to DNA constructs of the invention coding for transcription into tospovirus-RNA sequences of an open reading frame in viral complementary sense (i. e. having negative polarity), or into corresponding RNA sequences in which one or more codons have been replaced by their synonyms, or into RNA sequences homologous thereto.

A further preferred embodiment of the invention relates to DNA constructs of the invention coding for transcription into tospovirus-RNA sequences of a hairpin, or into RNA sequences homologous thereto.

Preferably the tospovirus-RNA sequence referred to hereinabove have at least 20, more preferably at least 50 nucleotides.

Examples of DNA constructs suitable for use according to the invention include TSWV Related DNA Sequences coding for transcription into (reference is made to Fig. 4 for S RNA nucleotide sequences, figure 6A for M RNA nucleotide sequences and figure 6B for L RNA nucleotide sequences):

i) the S RNA nucleotide sequence 1 to 2915;
ii) the S RNA nucleotide sequence 89 to 1483;
iii) the S RNA hairpin;
iv) the S RNA "pan-handle";
v) the S RNA nucleotide sequence 2763 to 1987;
vi) the L RNA nucleotide sequence 4462 to 1;
vii) the L RNA nucleotide sequence 41 to 2;
viii) the L RNA nucleotide sequence 3980 to 46;
ix) the L RNA nucleotide sequence (6706 + n) to (4462 + n); whereby n is the number of nucleotides of the gap between nucleotide 4462 (U) and the subsequently identified nucleotide (C) as shown in Fig. 6B.
x) the L RNA nucleotide sequence (6706 + n) to (6016 + n); wherein n is as defined hereinabove;
xi) the L RNA "pan-handle";
xii) an RNA sequence complementary to the S RNA nucleotide sequence 1987 to 2763;
xiii) an RNA sequence complementary to the S RNA nucleotide sequence 89 to 1483;
xv) an RNA sequence complementary to the L RNA nucleotide sequence 1 to 4462;
xv) an RNA sequence complementary to the L RNA nucleotide sequence 2 to 41;

xvi) an RNA sequence complementary to the L RNA nucleotide sequence 46 to 3980;
xvii) an RNA sequence complementary to the L RNA nucleotide sequence (4462 + n) to (6706 + n) wherein n is as defined above;
xviii) an RNA sequence complementary to the L RNA nucleotide sequence (6016 + n) to (6706 + n) wherein n is as defined above;
xix) S RNA nucleotide sequence 89 to 1483 in which one or more codons have been replaced by their synonyms;
xx) S RNA nucleotide sequence 2763 to 1987 in which one or more codons have been replaced by their synonyms;
xxi) L RNA nucleotide sequence 3980 to 236 in which one or more codons have been replaced by their synonyms;
xxii) L RNA nucleotide sequence 4462 to 236 in which one or more codons have been replaced by their synonyms;
xxiii) L RNA nucleotide sequence (6672 + n) to (4462 + n) in which one or more codons have been replaced by their synonyms;
xxiv) L RNA nucleotide sequence (6672 + n) to (6016 + n) in which one or more codons have been replaced by their synonyms;
xxv) the M RNA nucleotide sequence (m - 574) to m, wherein m is the total number of nucleotides of the M RNA;
xxvi) an RNA sequence complementary to the M RNA nucleotide sequence (m - 574) to m, wherein m is as defined hereinabove;
xxvii) RNA sequences homologous ot the nucleotide sequences defined under i) to xviii), xxv) or xxvi) hereinabove;
xxviii) the M RNA nucleotide sequence from (m - 28) to (m - 574) in which one or more of the codons have been replaced by their synonyms.
xxix) fragments of sequences defined under i) to xxviii) hereinabove.

Preferred TSWV Related DNA Sequences code for transcription into the RNA sequences according to ii), iii), iv), v), viii), ix) or xi) as defined above, or into RNA sequences homologous thereto, or into fragments thereof comprising at least 20, more preferably at least 50 nucleotides.

According to another preferred embodiment of the invention the DNA constructs of the invention comprise TSWV Related DNA Sequences coding for transcription into a combination of the 5' and 3' terminal sequences of the viral S, M or L RNA respectively, more preferably of the S or L RNA, in particular of the S RNA.

The invention further provides probes to diagnose suspected plants for infection with tospoviruses such as TSWV. Such probes comprise a labeled oligonucleotide (RNA or DNA) sequence complementary to an RNA sequence of tospoviruses such as TSWV (for the definition of the terms complementary see hereinbefore). The desired length of the sequence and appropriate method for diagnostic use of probes are known by those skilled in the art. A suitable probe will conveniently comprise a sequence of at least 15 preferably more than 30, more preferably from about 400 to 600 nucleotides complementary to an RNA sequence of tospoviruses such as TSWV.

Probes according to the invention are useful, in that they allow a person skilled in the art to identify tospovirus RNA or parts thereof in infected plant material, e.g. for diagnostic purposes prior to full expression of the disease symptoms.

The invention accordingly also provides a diagnostic method of determining tospovirus infection in plants which comprises detecting tospovirus replicative forms employing the probes of the invention to dot-blot type assays.

Probes according to the invention are useful, in that they allow a person skilled in the art to construct and to use chimeric genes comprising a DNA sequence corresponding to an RNA sequence of tospovirus.

The DNA constructs of the invention may be obtained by insertion of the TSWV Related DNA Sequence in an appropriate expression vector, such that it is brought under expression control of a promotor functioning in plants and its transcription terminated by a terminator.

The term appropriate expression vector as used herein refers to a vector containing a promoter and a terminator which function in plant cells.

The insertion of the TSWV Related DNA Sequence into an appropriate expression vector may be carried out in a manner known per se. Suitable procedures are also illustrated by the examples hereinafter.

Likewise the construction of appropriate expression vector may be carried out in a manner known per se.

The plants according to the invention may be obtained by

a) inserting into the genome of a plant cell a DNA construct according to the invention;
b) obtaining transformed cells; and
c) regenerating from the transformed cells genetically transformed plants.

The DNA vectors of the invention may be inserted into the plant genome of plants susceptible to TSWV infection. Such plant transformation may be carried out employing techniques known per se for the transformation of plants, such as the plant transformation techniques involving the Ti plasmids derived from *Agrobacterium tumefaciens, A. rhizogenes* or modifications thereof, naked DNA transformation or electroporation of isolated plant cells or organized plant structures, the use of micro-projectiles to deliver DNA, the use of laser systems, liposomes, or viruses or pollen as transformation vectors and the like.

The plants of the invention may be monitored for expression of a TSWV Related DNA Sequence by methods known in the art, including Northern analysis, Southern analysis, PCR techniques and/or immunological techniques. The plants of the invention show decreased susceptibility to TSWV infection as demonstrated by tests whereby said plants are exposed to TSWV preferentially at a concentration in the range where the rate of disease symptoms correlates linearly with the TSWV concentration in the inoculum.

Methods suitable for TSWV inoculation are known in the art; they include mechanical inoculation and, in particular, the use of appropriate vectors.

The plants of the invention may of course also be obtained by crossing of a obtained plant according to the methods of the invention with another plant to produce plants having in their plant genome a DNA construct of the invention.

The invention is illustrated by the following non-limitative examples and the attached figures.

Figure 1 gives an overview of the structure of the tomato spotted wilt virus.

Figure 2 gives a review of the cloning strategy employed for TSWV S RNA.

Figure 3 gives a review of the cloning strategy employed for TSWV L RNA.

Figure 4 shows the sequence and genomic organization of the TSWV S RNA. The amino acid sequence of the non-structural protein NSs (89-1483) is outlined above, that of the nucleocapsid protein N (2763-1987) under the corresponding RNA nucleotide sequence.

Figure 5 gives the distribution of translation initiation and termination codons for all reading frames of the S RNA (fig. 5A) and the L RNA (fig. 5B). Full bars indicate stop codons, half bars ATG start codons.

Figure 6 shows the nucleotide sequence of the TSWV M RNA (fig. 6A) and the L RNA (fig. 6B) as elucidated uptill now.

Figure 7 gives a schematic review of the construction of a suitable expression vector (pZU-A).

Figure 8 gives a schematic review of the construction of a suitable expression vector (pZU-B).

Figure 9 gives a schematic review of the construction of a suitable plasmid (pTSWV-N) comprising the nucleocapsid N protein gene.

Figure 10 gives a schematic review of the construction of a suitable plasmid (pTSWV-NSs) comprising the nucleocapsid NSs-protein gene.

Figure 11 gives a schematic example of the construction of a plant transformation vector pTSWV-NAB a DNA construct according to the invention.

Figure 12 gives a schematic example of the construction of a plant transformation vector pTSWV-Nmut BB, a DNA construct according to the invention.

Figure 13 gives a schematic example of the construction of a plant transformation vector pTSWV-NSsAB a DNA construct according to the invention.

Figure 14 gives a schematic example of the construction of a plant transformation vector pTSWV-NSsmut BB, a DNA construct according to the invention.

Figure 15 shows the "pan-handle" region of TSWV S RNA.

Figure 16 shows the hairpin region of TSWV S RNA.

Figure 17 shows a dot blot analysis of suspected plants.

Suitable examples of preferred TSWV Related DNA Sequences coding for transcription into a sequence of the hairpin structure of S RNA or of RNA sequences homologous thereto are sequences coding for the 1583-1708 nucleotide sequence of S RNA, for the 1709-1835 nucleotide sequence of S RNA for the 1583-1835 nucleotide sequence of the S RNA or for a sequence homologous to such sequences.

Suitable examples of preferred TSWV Related DNA Sequences coding for transcription into a sequence of the "pan-handle" region of S RNA or of RNA sequences homologous thereto is the combination of the sequences coding for the 1-70 and 2,850-2916 nucleotide sequence of S RNA or for sequences homologous to such sequences.

Suitable examples of preferred TSWV Related DNA Sequences coding for transcription into a "pan-handle" region of L RNA or of RNA sequences homologous thereto is the combination of sequences coding for the first 80, in particular the first 65 nucleotides from the 5'-end of the viral L RNA and the last 80, in

particular the last 65 nucleotides from the 3'-end of the viral L RNA, or for sequences homologous to such sequences or derivatives thereof

Other objects, features, advantages of the present invention will become apparent from the following examples.

References are abbreviated to the first authors name, full references are listed later.

Material and methods

All TSWV RNA derived sequences presented in here are depicted as DNA sequences for the sole purpose of uniformity only. A person skilled in the art appreciates that this is done for convenience only.

Cultivars of *Nicotiana tabacum* and *Petunia hybrida*, used in plant transformation studies, are grown under standard greenhouse conditions. Axenic explant material is grown on standard MS media (Murashige and Skoog, 1962) containing appropriate phytohormones and sucrose concentrations.

E. coli bacteria are grown on rotary shakers at 37 °C in standard LB-medium. *Agrobacterium tumefaciens* strains are grown at 28 °C in MinA medium supplemented with 0.1 % glucose (Ausubel *et al.*, 1987).

In all cloning procedures the *E. coli* strain JM83, (F⁻⁻ Δ(*lac-pro*), *ara*, *rpsL*, ϕ80, d*lac*ZM15) is used as a recipient for recombinant plasmids.

Binary vectors are conjugated to *Agrobacterium tumefaciens* strain LBA 4404, a strain containing the Ti-plasmid *vir* region, (Hoekema *et al.*, 1983) in standard triparental matings using the *E. coli* HB101, containing the plasmid pRK2013 as a helper strain (Figurski and Helinski, 1979). Appropriate *Agrobacterium tumefaciens* recipients are selected on media containing rifampicin (50 µg/ml) and kanamycine (50 µg/ml).

Cloning of fragments in the vectors pUC19 (Yanish-Perron *et al.*, 1985), pBluescript (Stratagene), pBIN19 (Bevan *et al.*, 1984) or derivatives, restriction enzyme analysis of DNA, transformation to *E. coli* recipient strains, isolation of plasmid DNA on small as well as large scale, nick-translation, *in vitro* transcription, DNA sequencing, Southern blotting and DNA gel electrophoresis are performed according to standard procedures (Maniatis *et al.*, 1982; Ausubel et al., 1987).

Examples

Example 1: Isolation of TSWV particles and the genetic material therein

Tomato spotted wilt virus strain CPNH1, a Brazilian isolate from tomato, is maintained on tomato by grafting. Virus is purified from systemically infected *Nicotiana rustica* leaves, after mechanical inoculation essentially as described by Tas *et al.* (1977). It is essential to maintain all material used in this isolation procedure at 4 °C. Twelve days after inoculation of 100 grams of infected leaves are harvested and ground 5 - 10 seconds at low speed setting in 5 volumes extraction buffer (0.1 M NaH_2PO_4, 0.01 M Na_2SO_3, pH 7) in a Waring blender. The suspension is filtered through cheesecloth and the filtrate is centrifuged for 10 minutes at 16,000 x *g*. The resulting pellet is resuspended in three volumes resuspensionbuffer (0.01 M NaH_2PO_4, 0.01 M Na_2SO_3, pH 7). The pellet is dissolved by stirring carefully at 4 °C. After centrifugation for 10 minutes at 12,500 x *g* the pellet is discarded and the supernatant is centrifuged again for 20 minutes at 50,000 x *g*. The pellet is resuspended in 0.2 volume of resuspensionbuffer (0.01 M NaH_2PO_4, 0.01 M Na_2SO_3, pH 7) and kept on ice for 30 minutes. Antiserum raised in rabbits against material from non-infected *Nicotiana rustica* is added to the solution, carefully stirred for 1 hour and non-viral complexes are pelleted by 10 minutes centrifugation at 16,000 x *g*. The cleared supernatant is loaded on a linear 5 - 40 % sucrose gradient in resuspensionbuffer: 0.01 M NaH_2PO_4, 0.01 M Na_2SO_3, pH 7, and spun for 45 minutes at 95,000 x *g*. The opalescent band containing TSWV virions is carefully collected with a syringe and diluted 4 times with resuspensionbuffer. Washed virions are pelleted by centrifugation for 1.5 hours at 21,000 x *g* and resuspended in one volume resuspensionbuffer. Generally, 100 grams of leaf material yields approximately 0,5 mg of TSWV virions. TSWV RNA is recovered preferentially from purified virus preparations by SDS-phenol extractions followed by ethanol precipitation. From 1 mg TSWV 1-5 µg of RNA is extracted. The intactness of the isolated RNA molecules is analysed by electrophoresis on an agarose gel. Three distinct RNA molecules are identified with apparent sizes of 2900 nucleotides (S RNA), 5000 nucleotides (M RNA) and 7500 nucleotides (L RNA) respectively.

Example 2: Sequence determination of the 3′-termini of the TSWV viral RNAs

For direct RNA sequencing TSWV RNA is extracted from purified nucleocapsids essentially according to Verkleij *et al* (1983) Twelve days after inoculation 100 grams of infected leaves are harvested and grounds 5 - 10 seconds at low speed setting in four volumes TAS-E buffer (0 01 M EDTA, 0 01 M Na_2SO_3, 0.1 % cysteine, 0 1 M TRIS pH 8.0) in a Waring blender. The suspension is filtered through cheesecloth and centrifuged 10 minutes at 1,100 x *g*. Nucleocapsids are recovered from this supernatant by 30 minutes centrifugation at 66,000 x *g*. The pellet is carefully resuspended in one volume TAS-R buffer (1 % Nonidet NP-40 0 01 M EDTA, 0 01 M Na_2SO_3, 0.1 % cysteine, 0.01 M glycine, 0 01 M TRIS pH 7 9) The pellet is dissolved by stirring carefully for 30 minutes at 4 °C. The supernatant is cleared by a 10 minutes centrifugation at 16,000 x *g*. Crude nucleocapsids are collected from the cleared supernatant by sedimentation through a 30 % sucrose cushion for 1 hour at 105,000 x *g*. The nucleocapsid pellet is resuspended in 400 µl 0 01 M Na-citrate pH 6 5, layered on a 20 - 40 % sucrose (in 0 01 M Na-citrate pH 6 5) and spun for 2 hours at 280,000 x *g*. The three different opalescent bands, respectively L, M and S nucleocapsid are collected separately. TSWV RNA is recovered preferentially from purified nucleocapsid preparations by SDS-phenol extractions followed by ethanol precipitation. Routinely 100 µg of RNA are obtained from 100 grams of infected leaves. The 3′-ends of the separate TSWV RNAs are labeled using RNA ligase and 5′-[^{32}P]pCp. These end-labeled RNA molecules are separated on a low gelling temperature agarose gel (Wieslander, 1979). The enzymatic approach described by Clerx-Van Haaster and Bishop (1980) and Clerx-Van Haaster *et al* (1982) is used to determine the 30 terminal nucleotides of the 3′- and 5′-ends of both S and M RNA.

Synthetic oligonucleotides complementary to the 3′-termini are synthesized using a commercially available system (Applied Biosystems) and used for dideoxy-sequencing with reverse transcriptase

Example 3: cDNA cloning of TSWV genetic material

Oligonucleotides complementary to the 3′-ends of the S and L RNA are used for priming first strand cDNA synthesis With these primers, double stranded DNA to TSWV RNA is synthesized in principle according to Gubler and Hoffman (1983) Two different approaches are used to generate cDNA clones to the TSWV viral RNAs A first series of clones is obtained by random priming of the TSWV RNA using fragmented single stranded calf thymus DNA, followed by first and second strand cDNA synthesis cDNA is made blunt-ended using T4-DNA polymerase and ligated with T4 ligase into the SmaI site of pUC19 A second series of TSWV cDNA clones is obtained by priming first strand DNA synthesis with the oligonucleotides complementary to the 20 terminal nucleotides at the 3′-ends of the TSWV RNAs. After second strand synthesis, treatment with T4 DNA polymerase to create blunt ends, phosphorylated EcoRI linkers are ligated to the ends of the cDNAs essentially according to Huynh *et al* (1985) After restriction of these cDNA molecules with EcoRI the cDNA fragments are ligated in the EcoRI site of lambda gt10 cDNA clones from both series containing viral inserts are selected via colony hybridization, essentially according to Grunstein and Hogness (1975) using [^{32}]P-labeled, randomly primed first strand cDNA as a probe. Sets of overlapping cDNA clones are selected by Southern analysis followed by plasmid and/or phage walking in order to construct restriction maps, based on cDNA derived sequences of the S (figure 2) and L RNA (figure 3)

Example 4: Sequence determination of the TSWV S M and L RNA

In order to determine the sequence of the S RNA 4 selected cDNA clones are subcloned in M13mp18 resulting in the plasmids pS614, pS608, pS520 and pS514, as described above (fig. 2). These clones are sequenced in both directions using the standard protocol of Yanish-Perron *et al* (1985) The nucleotide sequence of the 3′-end of the S RNA is determined by primer extension of the synthetic oligonucleotide S1 (5′ d(GAGCAATCGTGTCAATTTTG)), which is complementary to the 20 nucleotides of the 3′-terminus A second synthetic oligonucleotide S3 complementary to the nucleotides 30 to 50 from the 5′-end of the viral S RNA is used to verify the 5′-terminal sequence of the S RNA by primer extension. Sequence data from the TSWV SRNA (2916 nt) are summarized in figure 4

Computer simulated translation of the 6 different reading frames on the viral strand and viral complementary strand reveals the presence of two putative open reading frames (see figure 5A) On the viral strand an open reading frame is found starting at position 89 and terminating at an UAA stopcodon at

position 1483 possibly encoding a protein of 464 amino acids with a molecular weight of 52.4 kd. This protein is most likely a non-structural protein, tentatively designated NSs. The other open reading frame is located on the viral complementary strand from position 2763 to 1987, encoding a 258 amino acids long polypeptide with a molecular weight of 28.8 kd. This open reading frame possibly encodes the viral nucleocapsid protein N.

To verify this assumption the amino acid composition of the purified nucleocapsid protein is determined and compared with the deduced composition from the sequence data. Also, the putative nucleocapsid encoding N protein gene is inserted into pBluescript. Using T7 RNA polymerase the N protein gene is transcribed *in vitro*. This *in vitro* synthesized transcript is subsequently translated in an *in vitro* rabbit reticulocyte lysate (New England Nuclear: NEN) system in the presence of [^{35}S] labeled methionine (NEN). A protein with a molecular weight identical to the native nucleocapsid protein could be precipitated from the synthesized radioactive proteins, when antibodies raised against purified nucleocapsid protein are used, following procedures essentially as described (Van Grinsven *et al*, 1986). This indicates that the N protein gene encodes the viral nucleocapsid protein. Figure 5A shows the coding capacities of the viral and the viral complementary strand of the S RNA indicating the NSs and N protein gene respectively, both genes being expressed from subgenomic mRNAs. Thus, the unique situation occurs that a plant virus RNA has an ambisense gene arrangement. Other important features of this S RNA sequence is the existence of complementary terminal repeats capable of forming so-called "pan-handle" structures. These structures play an important role in replication and transcription of the viral RNA. Another putative regulatory element is the hairpin stucture in the intergenic region of the S RNA, which most likely contains the transcription termination signals for both subgenomic mRNAs, encoding respectively the N and NSs-protein.

The nucleotide sequence of the TSWV M and L RNA was elucidated employing the same strategies and methods followed to determine the nucleotide sequence of the S RNA. Figure 3 shows the cloning strategy for the L RNA derived, overlapping cDNA clones. Summarized sequence data of the TSWV M and L RNA are given in figure 6A and 6B respectively. Computer simulated translation suggests that this viral L RNA is of negative polarity containing one single open reading frame which starts 34 nucleotides from the 3'-end of the viral RNA and stops at nucleotide 236 from the 5'-end of the viral RNA (figure 5B). The mRNA (complementary to the viral L RNA) containing this open reading frame is most probably translated into the viral transcriptase.

Example 5: Construction of an expression vector pZU-A

The 35S cauliflower mosaic virus (CaMV) promoter fragment is isolated from the recombinant plasmid pZO27, a derivative of pUC19 carrying as a 444 bp HindIII-PstI fragment the HincII-HphI region of the 35S promoter of CaMV strain Cabb-S (Franck et al., 1980). The nucleotide sequences of CaMV strains are very similar for the different strains. The 35S promoter fragment is excised from pZO27 as a 472 bp EcoRI-PstI fragment which contains: a part of the polylinker region, 437 bp of the non-transcribed region, the transcription initiation site and 7 bp of the non-translated leader region but not containing any 35S translational initiators. This 35S promoter fragment is ligated using T4 ligase into EcoRI-PstI linearized pZO008. This plasmid pZO008 carries the nopaline synthase (NOS) terminator as a 270 bp PstI-HindIII fragment. The resulting recombinant plasmid pZU-A carries the 35S promoter, a unique PstI site and the NOS terminator (figure 7). The sequence of the used 35S promoter in the plant expression vector pZU-A is as follows:

```
   1 AAGCTTCTAG AGATCCGTCA ACATGGTGGA GCACGACACT CTCGTCTACT
  51 CCAAGAATAT CAAAGATACA GTCTCAGAAG ACCAAAGGGC TATTGAGACT
 101 TTTCAACAAA GGGTAATATC CGGAAACCTC CTCGGATTCC ATTGCCCAGC
 151 TATCTGTCAC TTCATCAAAA GGACAGTAGA AAAGGAAGGT GGCACCTACA
 201 AATGCCATCA TTGCGATAAA GGAAAGGCTA TCGTTCAAGA TGCCTCTGCC
 251 GACAGTGGTC CCAAAGATGG ACCCCCACCC ACGAGGAGCA TCGTGGAAAA
```

```
301 AGAAGACGTT CCAACCACGT CTTCAAAGCA AGTGGATTGA TGTGATATCT
351 CCACTGACGT AAGGGATGAC GCACAATCCC ACTATCCTTC GCAAGACCCT
401 TCCTCTATAT AAGGAAGTTC ATTTCATTGG AGAGGACCCT GCAG
```

The complete sequence from the used NOS terminator in the vector pZU-A is as follows:

```
  1 CTGCAGATCG TTCAAACATT TGGCAATAAA GTTTCTTAAG ATTGAATCCT
 51 GTTGCCGGTC TTGCGATGAT TATCATATAA TTTCTGTTGA ATTACGTTAA
101 GCATGTAATA ATTAACATGT AATGCATGAC GTTATTTATG AGATGGGTTT
151 TTATGATTAG AGTCCCGCAA TTATACATTT AATACGCGAT AGAAAACAAA
201 ATATAGCGCG CAACCTAGGA TAAATTATCG CGCGCGGTGT CATCTATGTT
251 ACTAGATCTC TAGAAAGCTT
```

Example 6: Construction of an expression vector pZU-B

The recombinant plasmid pZO347 is a derivative of pBluescript carrying a 496 bp BamHI-SmaI fragment containing a 426 bp 35S promoter fragment (HincII fragment) of CaMV strain Cabb-S, linked to a 67 bp fragment of the non-translated leader region, the so-called Ω-region, of the tobacco mosaic virus. This results in a chimeric promoter with a complete transcriptional fusion between the promoter of CaMV to the untranslated leader of TMV. By using *in vitro* mutagenesis the original position of the TMV ATG startcodon is mutated to a SmaI site.

The plasmid pZO008 carries the nopaline synthase (NOS) terminator as a 260 bp PstI-HindIII fragment. This PstI-HindIII fragment is excised from pZO008 and ligated using T4 ligase into PstI-HindIII linearized pZO347. The resulting recombinant plasmid pZU-B is another plant expression vector. The sequence of this 35S-Ω promoter as used in the plant expression vector pZU-B is as follows:

```
  1 GGATCCGGAA CATGGTGGAG CACGACACGC TTGTCTACTC CAAAAATATC
 51 AAAGATACAG TCTCAGAAGA CCAAAGGGCA ATTGAGACTT TTCAACAAAG
101 TTATTGTGAA GATAGTGGAA AAGGAAGGTG GCTCCTACAA ATGCCATCAT
151 TGCGATAAAG GAAAGGCCAT CGTTGAAGAT GCCTCTGCCG ACAGTGGTCC
201 CAAAGATGGA CCCCCACCCA CGAGGAGCAT CGTGGAAAAA GAAGACGTTC
251 CAACCACGTC TTCAAAGCAA GTGGATTGAT GTGATATCTC CACTGACGTA
301 AGGGATGACG CACAATCCCA CTATCCTTCG CAAGACCCTT CCTCTATATA

351 AGGAAGTTCA TTTCATTTGG AGAGGACTTT TTACAACAAT TACCAACAAC
401 AACAAACAAC AAACAACATT ACAATTACTA TTTACAATTA CCCGGG
```

The resulting recombinant plasmid pZU-B contains the 35S HincII-TMV Ω fusion (35S-Ω), unique SmaI and PstI sites and the NOS terminator (figure 8). This expression vector is preferentially used in constructing translational fusions of the gene to be expressed downstream of the chimeric promoter 35S-Ω.

Example 7: Subcloning of the TSWV-N protein gene

The TSWV-N protein coding sequence is obtained by fusion of the cDNA clones pS614 and pS520 (see figure 2). The cDNA clone pS520 is subjected to EcoRI-HindIII double-digestion and the fragment containing the 3′-end of the TSWV-N protein gene is separated electrophoretically and purified from the gel using a DEAE membrane (NA-45, Schleicher and Schüll) and cloned in the EcoRI-HindIII linearized pBluescript (Stratagene) resulting in the recombinant plasmid pS520E/H. The 5′-end containing fragment of the TSWV-N protein is excised from pS614 by a EcoRI digestion. This fragment is separated electrophoretically and purified from the gel using a DEAE membrane (NA-45, Schleicher and Schüll) and cloned in the EcoRI linearized pS520E/H resulting in the recombinant plasmid pTSWV-N3. The TSWV-N gene containing plasmid pTSWV-N3 is linearized by digestion with PvuII. PstI linkers 5′ d(CCTGCAGG) are ligated with T4 ligase to the blunt ends of the linear DNA. Subsequently, the DNA is digested with PstI and the fragment containing the TSWV-N protein sequence is separated electrophoretically and excised from the gel. The TSWV-N protein gene containing fragment is ligated into a PstI linearized vector such as pBluescipt (Stratagene) to yield the recombinant plasmid pTSWV-N (figure 9). This addition of restriction sites facilitates the construction of further plasmids. (Alternatively, one may choose to add the sites in different ways such as but not limited to site-directed mutagenesis or by ligation of other synthetic oligonucleotide linkers. These methods are all known to a person skilled in the art.)

Example 8: Subcloning of the TSWV non-structural protein gene (NSs-gene) of the TSWV S RNA

The sequence of the non-structural protein NSs is isolated from the cDNA clone pS514. The NSs-protein gene is located on an EcoRI fragment. After restriction of the cDNA clone pS514 with EcoRI and treatment with T4 DNA polymerase to create blunt ends. The NSs gene containing fragment is separated electrophoretically on an agarose gel and excised from the gel. To this blunt ended fragment containing the NSs-protein gene synthetic PstI linkers (5′ d(CCTGCAGG)-3′) are ligated using T4 polymerase. After restriction with PstI, the NSs-protein gene containing fragment is ligated in an PstI linearized pBluescript to yield the recombinant plasmid pTSWV-NSs (figure 10).

Example 9: Construction of plant transformation vectors containing TSWV sequences

Example 9A: N-protein gene constructions in pZU-A

In order to create a plant transformation vector containing the N protein gene driven by the 35S promoter and terminated by the NOS terminator, the PstI fragment of pTSWV-N is isolated and inserted into PstI linearized pZU-A thereby creating the chimeric gene cassette vector pTSWV-NA. The cassette containing the 35S promoter, the N-protein gene and the NOS terminator is excised from pTSWV-NA by restriction with XbaI and ligated in the unique XbaI site of pBIN19, a binary transformation vector developed by Bevan *et al* (1984). The resulting plasmid pTSWV-NAB (figure 11) is used in plant transformation experiments using methods well known to a person skilled in the art.

Example 9B: N-protein constructions in pZU-B

In order to make a fusion in which the ATG start codon from the N protein gene is fused directly to the 3′-end of the TMV untranslated leader of the 35S-Ω promoter the startcodon of the N gene has to be mutated. Using site-directed mutagenesis, the sequence 5′ d(ACGATCATC**ATG** TCT) in pTSWV-N is mutated to 5′ d(ACGATATC**ATG** TCT), thereby creating an EcoRV site: 5′ d(GAT|ATC) just proximal to the ATG startcodon of the N gene. The resulting recombinant plasmid is called pTSWV-Nmut. The mutated N protein gene is excised from this plasmid via an EcoRV-PstI digestion. This fragment is isolated and inserted into the SmaI-PstI linearized pZU-B, resulting in recombinant plasmid pTSWV-NmutB. The chimeric cassette containing the 35S-Ω promoter, the mutated N gene and the NOS terminator is excised from the plasmid pTSWV-NmutB via a BamHI/XbaI digestion. The isolated chimeric gene cassette is then inserted into the BamHI/XbaI linearized pBIN19 to create the binary transformation vector pTSWV-NmutBB. The resulting plasmid pTSWV-NmutBB (figure 12) is used in plant transformation experiments using methods well known to a person skilled in the art.

Example 9C: NSs-protein gene constructions in pZU-A

In order to create a plant transformation vector containing the NSs-protein gene driven by the 355 promoter and terminated by the NOS terminator, the PstI fragment of pTSWV-NSs is isolated and inserted into PstI linearized pZU-B to create the chimeric gene cassette vector pTSWV-NsA. The cassette containing the 35S promoter, the NSs-protein gene and the NOS terminator is excised from pTSWV-NsA by restriction with EcoRI and XbaI and ligated into EcoRI-XbaI linearized pBIN19. The resulting plasmid pTSWV-NsAB (figure 13) is used in plant transformation experiments using methods well known to a person skilled in the art

Example 9 D: NSs-protein gene constructions in pZU-B

In order to create a fusion in which the ATG start codon from the NSs-protein is fused directly to the 3′-end of the TMV leader of the 35S-Ω promoter the startcodon of the NSs gene has to be mutated. Using site-directed mutagenesis a procedure known to a person skilled in the art, the sequence 5′ d-(AACCATA**ATG** TCT) is mutated to 5′ d(AACCAT**ATG** TCT), thereby creating a NdeI site: 5′ d(CA|TATG) that includes the ATG startcodon of the NSs-protein gene. The resulting plasmid is called pTSWV-NSsmut. The plasmid pTSWV-NSsmut is linearized with NdeI, followed by a treatrnent with T4 DNA polymerase to create blunt ends. This linearized DNA is digested with PstI and the fragment containing the mutated NSs gene is isolated and inserted into SmaI PstI linearized pZU-B resulting in the recombinant plasmid pTSWV-NsmutB. The chimeric cassette containing the 35S-Ω promoter, the mutated NSs-protein gene and the NOS terminator is excised from the plasmid pTSWV-NsmutB via a BamHI/XbaI digestion. The isolated chimeric gene cassette is then inserted into the BamHI/XbaI linearized pBIN19 to create the binary transformation vector pTSWV-NsmutBB. The resulting plasmid pTSWV-NsmutBB (figure 14) is used in plant transformation experiments using methods well known to a person skilled in the art

Example 9E: 5′- and 3′-termini "pan-handle" constructions in pZU-A and pZU-B

A DNA analysis programme is used to locate the "pan-handle" loop in the viral TSWV S RNA. The strongest "pan-handle" loop that is detected includes the first 70 nucleotides at the 5′-end (1 to 70) of the viral S RNA and the last 67 nucleotides at the 3′-end (2850 to 2916) of the viral S RNA (figure 15). The DNA sequence containing this "pan-handle" loop in the viral S RNA is as follows:

```
   1 AGAGCAATTG TGTCAGAATT TTGTTCATAA TCAAACCTCA CTTAGAAAAT
  51 CACAATACTG TAATAAGAAC
```

and

```
2850 GTTCTTAATG TGATGATTTG TAAGACTGAG TGTTAAGGTA TGAACACAAA
2900 ATTGACACGA TTGCTCT
```

A DNA analysis programme is used to locate the "pan-handle" loop in the viral TSWV L RNA. A strong "pan-handle" loop that is detected includes the first 80 nucleotides at the 5′-end (1 to 80) of the viral L RNA and the last 80 nucleotides at the 3′-end of the viral L RNA. The DNA sequence containing this "pan-handle" loop in the viral L RNA is as follows:

```
  1 AGAGCAATCA GGTACAACTA AAACATATAA CCTCTCCACA GCCAGACTTT
 51 ACAAATTACA TAAGAATTCC CTCCAGTGAA
```

and

```
AAAGTGGTTC CATTTTCTAT TAATTTTTGT ATTTTCTGGA TGTTCATGTT
TGCTTAAAAT CGTTGTTACC TGATTGCTCT
```

These regions are synthesized on a commercial DNA synthesizer and appropriate linker sequences are added. Construction of the "pan-handle" vectors of S and L RNA results in respectively: pTSWV-termS and pTSWV-termL. Using appropriate restriction enzyme combination these fragments are inserted between the CaMV 35S promoter and the NOS terminator of pZU-A or between the 35S-Ω promoter and the NOS terminator of pZU-B yielding the chimeric cassettes: pTSWV-termSA, pTSWV-termLA, pTSWV-termSB and pTSWV-termLB. These cassettes are then transferred into the binary transformation vector pBIN19 using appropriate enzyme combinations yielding the following plasmids: pTSWV-termSAB, pTSWV-termLAB, pTSWV-termSBB and pTSWV-termLBB. Alternatively, it is possible to design "pan-handle" constucts including the 3′- and 5′-end termini that are larger as indicated above, or separated by any other DNA sequence in order to enhance the stability of the transcripts produced from these recombinant genes in plants. All "pan-handle" constructs resemble shortened tospovirus RNA, respectively TSWV RNA molecules and therefore can be regarded as defective interfering RNAs

Example 9F: Construction containing TSWV S RNA hairpin region in pZU-A and pZU-B

A DNA analysis programme is used to locate the hairpin loop in the viral TSWV S RNA. The strongest hairpin loop that is detected starts at nucleotide 1592 and ends at nucleotide 1834 (figure 16). The sequence containing this hairpin loop is as follows:

```
   1 TAGTAGAAAC CATAAAAACA AAAAATAAAA ATGAAAATAA AATTAAAATA
  51 AAATAAAATC AAAAAATGAA ATAAAAACAA CAAAAAATTA AAAAACGAAA
 101 AACCAAAAAG ACCCGAAAGG GACCAATTTG GCCAAATTTG GGTTTTGTTT
 151 TTGTTTTTTG TTTTTTGTTT TTTATTTTTT ATTTTATTTT TATTTTATTT
 201 TATTTTTATT TTATTTTTAT TTTATTTATT TTTTGTTTTC GTTGTTTTTG
 251 TTA
```

A HindIII fragment of 526 bp carrying the hairpin region is isolated from pS514. This fragment is excised from a agarose gel and subsequently treated with T4 polymerase to create blunt ends. In the following step PstI linkers are ligated to these blunt ends. After digestion with PstI the fragment is cloned in PstI linearized pZU-A, resulting in the recombinant plasmid pTSWV-HpSA. The plasmid pTSWV-HpSA is digested with HindIII and the fragment containing the chimeric gene is excised from an agarose gel and ligated into HindIII linearized pBIN19, resulting in the transformation vector pTSWV-HpSAB.

Alternatively, the HindIII fragment of pS514 is treated with T4 DNA polymerase to create blunt ends and is subsequently cloned in the SmaI site of the expression vector pZU-B, resulting in the recombinant plasmid pTSWV-HpSB. The plasmid pTSWV-HpSB is digested with HindIII and the fragment containing the chimeric gene is excised from an agarose gel and ligated into XbaI linearized pBIN19, resulting in the transformation vector pTSWV-HpSBB.

(It is clear to a person skilled in the art that also other fragments can be isolated from the cDNA clones of the TSWV S RNA containing the hairpin region as described above without interference with the function. Also, a fragment containing the hairpin region may be synthesized using a DNA-synthesizer.)

Example 10: Transformation of binary vectors to plant material

Methods to transfer binary vectors to plant material are well established and known to a person skilled in the art Variations in procedures exist due to for instance differences in used *Agrobacterium* strains different sources of explant material differences in regeneration systems depending on as well the cultivar as the plant species used

The binary plant transformation vectors as described above are used in plant transformation experiments according to the following procedures The constructed binary vector is transferred by tri-parental mating to an acceptor *Agrobacterium tumefaciens* strain, followed by southern analysis of the exconjugants for verification of proper transfer of the construct to the acceptor strain, inoculation and cocultivation of axenic explant material with the *Agrobacterium tumefaciens* strain of choice, selective killing of the *Agrobacterium tumefaciens* strain used with appropriate antibiotics, selection of transformed cells by growing on selective media containing kanamycine, transfer of tissue to shoot-inducing media transfer of selected shoots to root inducing media, transfer of plantlets to soil, assaying for intactness of the construct by southern analyses of isolated total DNA from the transgenic plant, assaying for proper function of the inserted chimeric gene by northern analysis and/or enzyme assays and western blot analysis of proteins

Example 11: Expression of TSWV RNA sequences in plant cells

RNA is extracted from leaves of regenerated plants using the following protocol Grind 200 mg leafmaterial to a fine powder in liquid nitrogen Add 800 µl RNA extraction buffer (100 mM Tris-HCl (pH 8,0) 500 mM NaCl, 2 mM EDTA, 200 mM β-Mercapto-ethanol. 0,4% SDS) and extract the homogenate with phenol, collect the nucleic acids by alcohol precipitation Resuspend the nucleic acids in 0 5 ml 10 mM Tris-HCl (pH 8,0), 1 mM EDTA add LiCl to a final concentration of 2 M, leave on ice for maximal 4 hours and collect the RNA by centrifugation Resuspend in 400 µl 10 mM Tris-HCl (pH 8,0), 1 mM EDTA and precipitate with alcohol, finally resuspend in 50 µl 10 mM Tris-HCl (pH 8,0), 1 mM EDTA RNAs are separated on glyoxal/agarose gels and blotted to Genescreen as described by van Grinsven *et al* (1986) TSWV viral RNA is detected using DNA or RNA probes labeled with [^{32}P], [^{35}S] or by using non-radioactive labeling techniques Based on these northern analysis, it is determined to what extent the regenerated plants express the chimeric TSWV genes

Plants transformed with chimeric constructs contalning a TSWV protein gene are also subjected to western blot analysis Proteins are extracted from leaves of transformed plants by grinding in sample buffer according to Laemmli (1970) A 50 µg portion of protein is subjected to electrophoresis in a 12,5 % SDS-polyacrylamide gel essentially as described by Laemmli (1970) Separated proteins are transferred to nitrocellulose electrophoretically as described by Towbin *et al* (1979). Transferred proteins are reacted with antiserum raised against purified TSWV nucleocapsids according to Towbin *et al* (1979) Based on the results of the western analysis, it is determined that transformed plants do contain TSWV proteins encoded by the inserted chimeric genes

Example 12: Resistance of plants against TSWV infections

Transformed plants are grown in the greenhouse under standard quarantine conditions in order to prevent any infections by pathogens. The transformants are self-pollinated and the seeds harvested Progeny plants are analyzed for segregation of the inserted gene and subsequently infected with TSWV by mechanical inoculation Tissue from plants systemically infected with TSWV is ground in 5 volumes of ice-cold inoculation buffer (10 mM phosphate buffer supplemented with 1% Na_2SO_3) and rubbed in the presence of carborundum powder on the first two fully extended leafs of approximately 5 weeks old seedlings Inoculated plants are monitored for symptom development during 3 weeks after inoculation

Plants containing TSWV Related DNA Sequences show reduced susceptibility to TSWV infection as examplified by a delay in symptom development, whereas untransformed control plants show severe systemic TSWV symptoms within 7 days after inoculation.

Example 13: Use of synthetic oligonucleotides for diagnostic purposes

RNA is extracted from leaves of suspected plants using the following protocol: grind 1 gram of leaf material, preferentially showing disease symptoms, in 3 ml 100 mM Tris-HCl, 50 mM EDTA, 1 5 M NaCl and 2% CTAB (pH 8.0). After grinding, 1 ml of the homogenate is subjected to chloroform extraction and incubated at 65 °C for 10 minutes The inorganic phase is then collected and extracted with phenol/chloroform (1:1), followed by a last extraction with chloroform. The ribonucleic acids are isolated from the inorganic phase, containing the total nucleic acids, by adding LiCl to a final concentration of 2 M The preparation is incubated at 4 °C for 1 hour, after which the ribonucleic acids are collected by centrifugation. The ribonucleic acid pellet is resuspended in 25 µl 10 mM Tris-HCl, 1 mM EDTA (pH 8.0) The ribonucleic acids are recovered by standard alcohol precipitation The ribonucleic acid pellet is resuspended in 25 µl 10 mM Tris-HCl, 1 mM EDTA (pH 8 0).

1 µl of the purified ribonucleic acids is spotted on a nylon blotting membrane (e.g Hybond-N Amersham UK). The presence of TSWV in the plant is detected by standard hybridization, using any part or parts of the sequence isolated from virions or preferentially by designing synthetic oligomers on the basis of disclosed sequence information as a probe. (Alternatively, *in vitro* transcripts of regions of the TSWV genome are used to detect TSWV Related RNA Sequences in diseased plants) A diseased plant is diagnosed by the occurrence of hybridization at the dot containing RNA material from the diseased plant

Using procedures as described herein RNA is isolated from 12 pepper plants (A1 - D3) suspected to be virus infected selected from a field. Similarly RNA is extracted from 3 TSWV inoculated tobacco plants (E1 - E3) and from 3 non-inoculated tobacco plants (F1 -F3). From each plant an RNA sample of 1 µl is spotted onto a Hybond membrane. This filter is hybridized under standard conditions with an *in vitro* transcript synthesized from the cDNA clone pS614 using T3 polymerase and α-[^{32}P]UTP as an radioactive label Control non-infected plants (F1 - F3) do not show a signal, control TSWV infected plants (E1 - E3) do show a strong signal, whereas suspected pepper plants all show signals ranging in intensity from weak to strong (see figure 17)

References

Abel PP, Nelson RS, De B Hoffmann N, Rogers SG, Fraley RT and Beachy RN (1986) Delay of disease development in transgenic plants that express the tobacco mosaic virus coat protein gene Science 232, 738-743.

Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K (1987) Current protocols in molecular biology. Green Publishing Associates and Wiley Intersciences, New York, Chichester, Brisbane, Toronto and Singapore.

Beachy RN, Fraley RT and Rogers SG, EP 0 223 452: Protection of plants against viral infection Monsanto Company & Washington University

Bevan M (1984) Binary *Agrobacterium* vectors for plant transformation Nucl Acids Res 12, 8711-8721

Clerx-Van Haaster CM and Bishop DHL (1980) Analysis of the 3′-terminal RNA sequences of Snowshoe hare and Lacrosse Bunyaviruses Virolology 105, 564-574.

Clerx-Van Haaster CM, Clerx JPM, Ushijima H, Akashi H, Fuller F and Bishop DHL (1982) Analysis of the 3′-terminal RNA sequences of Bunyaviruses and Nairoviruses (Bunyaviridae): evidence of end sequence generic differences within the virus family J Gen Virol 61, 289-292.

Figurski D, Helinski DR (1979) Replication of an origin containing a derivative of plasmid RK_2 dependent on a plasmid function provided in *trans* Proc Natl Acad Sci USA 76: 1648-1652

Franck A, Guilley H, Jonard G, Richards K and Hirth L (1980) Cell 21, 285-294.

Grunstein JM and Hogness DS (1975) Colony hybridization: a method for the isolation of cloned DNAs that contain a specific gene Proc Natl Acad Sci USA 72, 3961-3965

Gubler U, Hoffman BJ (11983) A simple and very efficient method for generating cDNA libraries Gene 25, 263-269.

Hoekema A, Hirsch PR, Hooykaas PJJ, Schilperoort RA (1983) A binary vector system based on separation of vir- and T-region of the *Agrobacterium tumefaciens* Ti-plasmid. Nature 303: 179-180

Huynh TH, Young RA and Davis RW (1985) Construction and screening cDNA libraries in lambda gt10 and lambda gt11 In: DNA cloning techniques: a practical approach. (Glover D, ed.) Press IRL Oxford, pp 49-78.

Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4 Nature 244, 29-30.

Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning, a laboratory manual Cold Spring Harbor Laboratory, New York

Murashige T. Skoog F (1962) A revised medium for rapid growth and bio-assays with tobacco tissue cultures. Physiol Plant 15: 473-497

Tas PWL, Boerjan ML and Peters D (1977) The structural proteins of tomato spotted wilt virus. J Gen Virol 36, 81-91.

Towbin H, Staehelin T and Gordon J (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose: procedure and some applications. Proc Natl Acad Sci USA 76, 4350-4354.

Van Grinsven MQJM, Gielen JJL, Zethof JLA, Nijkamp HJJ and Kool AJ (1986) Transcriptional and post-transcriptional regulation of chloroplast gene expression in *Petunia hybrida*. Theor Appl Gen 73, 94-101

Verkleij FN and Peters D (1983) Characterization of a defective form of tomato spotted wilt virus. J Gen Virol 64, 677-686

Wieslander R (1979) A simple method to recover intact high molecular weight RNA and DNA after electrophoretic separation in low gelling temperature agarose gels. Anal Biochem 98, 305-309

Yanish-Perron C; Vieira J and Messing J (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequence of the M13mp18 and pUC19 vectors. Gene 33, 103-119.

Claims

1. Recombinant DNA constructs comprising a DNA sequence coding for transcription into
a) an RNA sequence of tospoviruses or an RNA sequence homologous thereto;
b) an RNA sequence according to a) encoding for a tospovirus protein in which one or more codons have been replaced by their synonyms (i.e. codons corresponding to the same amino acid or termination signal) or a part thereof; or an RNA sequence homologous there to, or
c) an RNA sequence complementary to an RNA sequence according to a) or b),
which DNA is under expression control of a promoter and a terminator functioning in plants

2. The DNA constructs of Claim 1, wherein the DNA sequences code for transcription into:
i) the S RNA nucleotide sequence 1 to 2915;
ii) the S RNA nucleotide sequence 89 to 1483;
iii) the S RNA hairpin;
iv) the S RNA "pan-handle";
v) the S RNA nucleotide sequence 2763 to 1987;
vi) L RNA nucleotide sequence 4462 to 1;
vii) the L RNA nucleotide sequence 41 to 2;
viii) the L RNA nucleotide sequence 3980 to 46;
ix) the L RNA nucleotide sequence (6706 + n) to (4462 + n); whereby n is the number of nucleotides of the gap between nucleotide 4462 (U) and the subsequently identified nucleotide (C) as shown in Fig. 6B
x) the L RNA nucleotide sequence (6706 + n) to (6016 + n); wherein n is as defined hereinabove;
xi) the L RNA "pan-handle";
xii) an RNA sequence complementary to the S RNA nucleotide sequence 1987 to 2763;
xiii) an RNA sequence complementary to the S RNA nucleotide sequence 89 to 1483;
xiv) an RNA sequence complementary to the L RNA nucleotide sequence 1 to 4462;
xv) an RNA sequence complementary to the L RNA nucleotide sequence 2 to 41;
xvi) an RNA sequence complementary to the L RNA nucleotide sequence 46 to 3980;
xvii) an sequence complementary to the L RNA nucleotide sequence (4462 + n) to (6706 + n), wherein n is as defined above;
xviii) an RNA sequence complementary to the L RNA nucleotide sequence (6016 + n) to (6706 + n) wherein n is as defined above;
xiv) S RNA nucleotide sequence 89 to 1483 in which one or more codons have been replaced by their synonyms;
xx) S RNA nucleotide sequence 2763 to 1987 in which one or more codons have been replaced by their synonyms;
xxi) L RNA nucleotide sequence 3980 to 236 in which one or more codons have been replaced by their synonyms;
xxii) L RNA nucleotide sequence 4462 to 236 in which one or more codons have been replaced by their synonyms;
xxiii) L RNA nucleotide sequence (6672 + n) to (4462 + n) in which one or more codons have been replaced by their synonyms, wherein n is as defined above;
xxiv) L RNA nucleotide sequence (6672 + n) to (6016 + n) in which one or more codons have been

replaced by their synonyms, wherein n is as defined above;
xxv) the M RNA nucleotide sequence (m-574) to m, wherein m is the total number of nucleotides of the M RNA;
xxvi) an RNA sequence complementary to the M RNA nucleotide sequence (m-574) to m wherein m is as defined hereinabove;
xxvii) RNA sequences homologous to the nucleotide sequences defined under i) to xviii), xxv) or xxvi) hereinabove;
xxviii) the M RNA nucleotide sequence from (m-26) to (m-574) in which one or more of the codons have been replaced by their synonyms, wherein m is as defined above;
xxix) a fragment of a DNA sequence defined under i) to xxviii) hereinabove

3. The DNA construct of Claim 1, coding for transcription in tospovirus-RNA sequences of a pan-handle, or into RNA sequences homologous thereto.

4. The DNA construct of Claim 1, coding for transcription into tospovirus-RNA sequences of an open reading frame in viral complementary sense, or into corresponding RNA sequences in which one or more codons have been replaced by their synonyms, or into RNA sequences homologous thereto.

5. The DNA construct of Claim 1, coding for transcription into tospovirus-RNA sequences of a hairpin, or into RNA sequences homologous there to.

6. The DNA construct of Claims 1 to 5, coding for transcription into tospovirus-RNA sequences, or into tospovirus-RNA sequences in which one or more codons have been replaced by their synonyms, or into RNA sequences homologous thereto of at least 20 nucleotides.

7. The DNA construct of Claim 6, coding for transcription into tospovirus-RNA sequences, or into tospovirus-RNA sequences in which one or more codons have been replaced by their synonyms, or into RNA sequences homologous thereto of at least 50 nucleotides.

8. The DNA construct of Claim 1, wherein the DNA sequence codes for transcription into a combination of the 5′n and 3′ terminal sequences of the viral S or L RNA respectively.

9. The DNA construct of Claims 1 to 8, wherein the promoter is a viral, fungal, bacterial, animal or plant derived promoter functioning in plant cells.

10. The DNA construct of Claim 9, wherein the terminator is a viral, fungal, bacterial, animal or plant derived terminator functioning in plant cells.

11. A plant comprising in its genome a DNA construct in accordance with Claims 1 to 10.

12. A probe comprising a single or double stranded oligonucleotide sequence complementary to an RNA sequence of a tospovirus.

13. A probe according to Claim 12, wherein the oligonucleotides are complementary to the tospovirus S RNA sequence, the tospovirus L RNA nucleotide sequence or to fragments of such sequences comprising at least 15 nucleotides.

14. A probe according to Claims 12 or 13 wherein the oligonucleotide sequence has from 400 to 600 nucleotides.

15. A process of preparing plants according to Claim 11, which comprises
a) inserting into the genome of plant cell a DNA construct of Claim 1,
b) obtaining transformed cells;
c) regeneration from the transformed cells genetically transformed plants

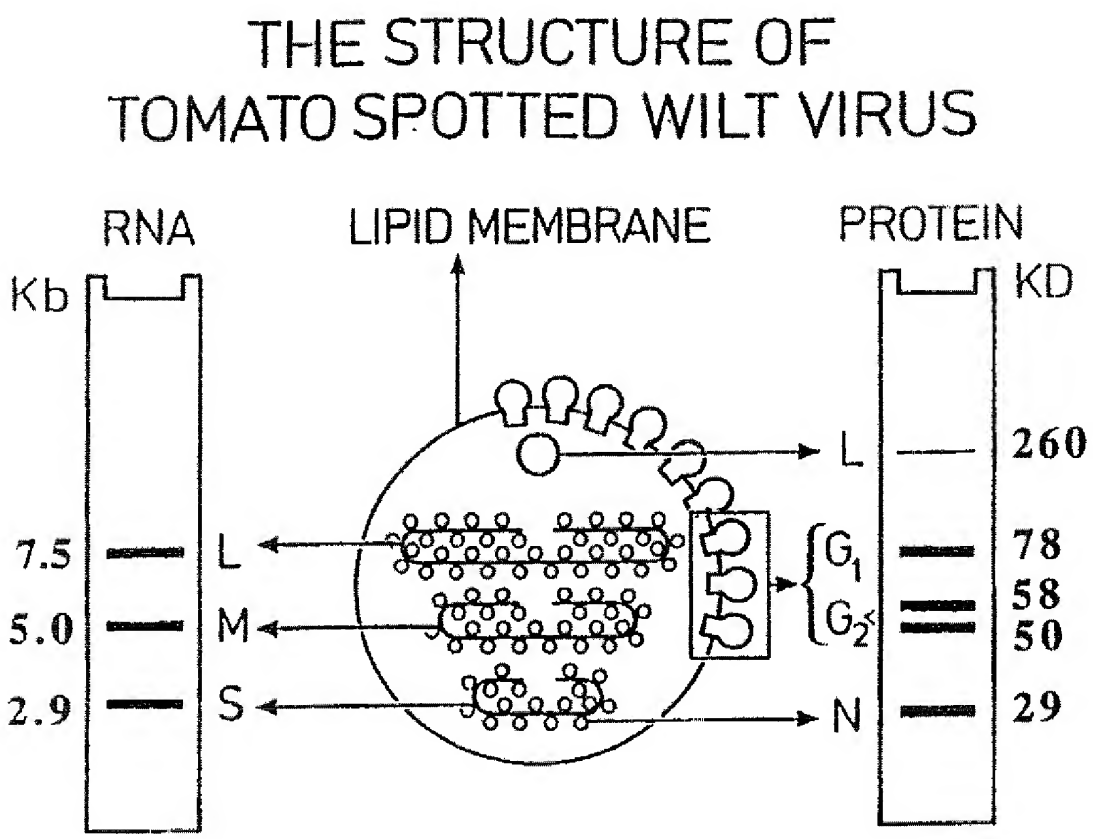
THE STRUCTURE OF
TOMATO SPOTTED WILT VIRUS
RNA
LIPID MEMBRANE
PROTEIN
Kb
KD
7.5
L
5.0
M
2.9
S
L
260
G1
78
58
G2
50
N
29

Figure 1

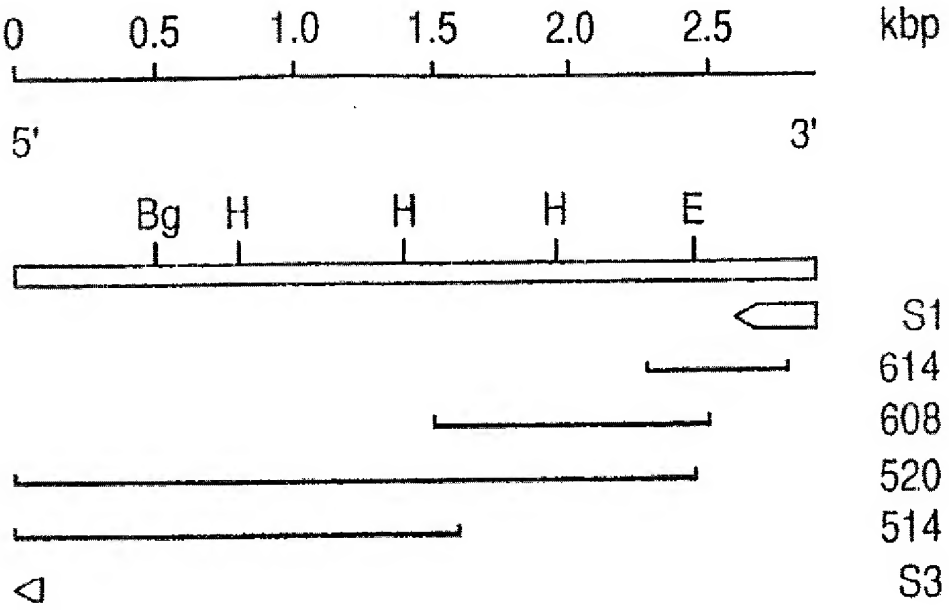

Figure 2

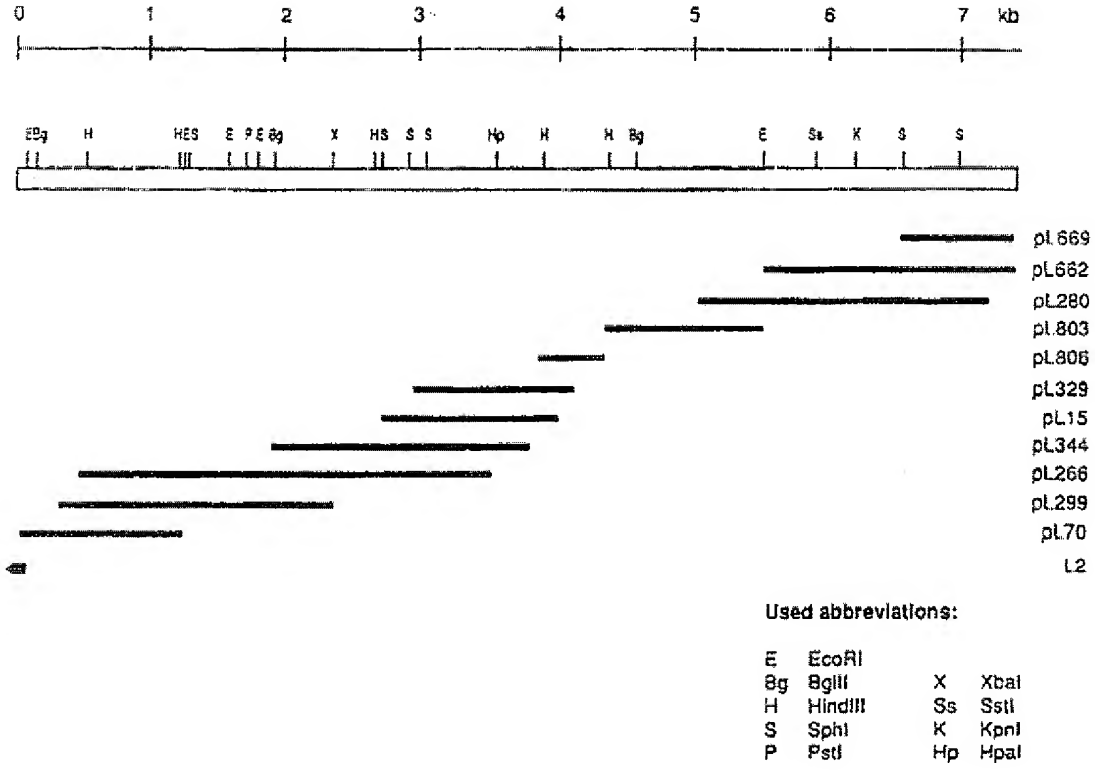
0
1
2
3
4
5
6
7
kb
pL669
pL662
pL280
pL803
pL806
pL329
pL15
pL344
pL266
pL299
pL70
L2
Used abbreviations:
E EcoRI
Bg BglII
H HindIII
S SphI
P PstI
X XbaI
Ss SstI
K KpnI
Hp HpaI

Figure 3

```
1 AGA GCA AUU GUG UCA GAA UUU UGU UCA UAA UCA AAC CUC ACU UAG AAA AUC ACA AUA CUG

                                                    met ser ser ser val tyr glu ser ile ile gln
61 UAA UAA GAA CAC AGU ACC AAU AAC CAU AAU GUC UUC AAG UGU UUA UGA GUC GAU CAU UCA

    thr arg ala ser val trp gly ser thr ala ser gly lys ala val val asp ser tyr trp
121 GAC AAG AGC UUC AGU CUG GGG AUC AAC UGC AUC UGG UAA AGC UGU UGU AGA UUC UUA CUG

    ile his glu leu gly thr gly ser gln leu val gln thr gln leu tyr ser asp ser arg
181 GAU UCA UGA ACU UGG UAC UGG UUC UCA ACU AGU UCA GAC CCA GCU GUA UUC UGA UUC AAG

    ser lys val val leu trp leu tyr cys lys val gly ile phe pro val lys lys lys arg
241 AAG CAA AGU AGU CCU UUG GCU AUA CUG CAA AGU AGG GAU CUU CCC UGU GAA GAA GAA GAG

    phe leu ser gln his val tyr ile pro ile phe asp asp ile asp phe ser ile asn ile
301 AUU UCU UUC UCA GCA UGU GUA UAU CCC UAU UUU UGA UGA UAU UGA UUU UAG CAU CAA UAU

    asp asn ser val leu ala leu ser val cys ser asn thr val asn ala asn gly val lys
361 UGA UAA CUC UGU UCU GGC ACU AUC UGU UUG CUC AAA UAC AGU CAA UGC UAA CGG AGU GAA

    his gln gly his leu lys val leu ser pro ala gln leu his ser ile glu ser ile met
421 ACA UCA AGG UCA UUU GAA GGU UUU GUC UCC UGC CCA GCU CCA CUC UAU UGA AUC UAU CAU

    asn arg ser asp ile thr asp arg phe gln leu gln glu lys asp ile ile pro asn asp
481 GAA CAG AUC UGA UAU UAC AGA CCG AUU CCA GCU CCA AGA AAA AGA CAU AAU UCC CAA UGA

    lys tyr ile glu ala ala asn lys gly ser leu ser cys val lys glu his thr tyr lys
541 CAA AUA CAU UGA AGC UGC AAA CAA AGG CUC UUU GUC UUG UGU CAA AGA GCA UAC CUA UAA

    ile glu met cys tyr asn gln ala leu gly lys val asn val leu ser pro asn arg asn
601 GAU CGA GAU GUG CUA UAA UCA GGC UUU AGG CAA AGU GAA UGU UCU AUC UCC UAA CAG AAA

    val his glu trp leu tyr ser phe lys pro asn phe asn gln val glu ser asn asn arg
661 UGU CCA UGA AUG GCU GUA CAG UUU CAA GCC AAA UUU CAA UCA AGU UGA AAG CAA CAA CAG

    thr val asn ser leu ala val lys ser leu leu met ser ala glu asn asn ile met pro
721 AAC UGU AAA UUC UCU UGC AGU GAA AUC UCU GCU CAU GUC AGC AGA AAA CAA CAU CAU GCC

    asn ser gln ala ser thr asp ser his phe lys leu ser leu trp leu arg val pro lys
781 UAA CUC UCA AGC UUC CAC UGA UUC UCA UUU CAA GCU GAG CCU CUG GCU AAG GGU UCC AAA

    val leu lys gln val ser ile gln lys leu phe lys val ala gly asp glu thr asn lys
841 GGU UUU GAA GCA GGU UUC CAU UCA GAA AUU GUU CAA GGU UGC AGG AGA UGA AAC AAA CAA

    thr phe tyr leu ser ile ala cys ile pro asn his asn ser val glu thr ala leu asn
901 AAC AUU UUA UUU AUC UAU UGC CUG CAU UCC AAA CCA UAA CAG UGU UGA GAC AGC UUU AAA

    ile thr val ile cys lys his gln leu pro ile arg lys cys lys ala pro phe glu leu
961 CAU UAC UGU UAU UUG CAA GCA UCA GCU CCC AAU UCG CAA AUG CAA AGC UCC UUU UGA AUU

    ser met met phe ser asp leu lys glu pro tyr asn ile val his asp pro ser tyr pro
1021 AUC AAU GAU GUU UUC UGA UUU AAA GGA GCC UUA CAA CAU UGU UCA UGA CCC UUC AUA CCC

    lys gly ser val pro met leu trp leu glu thr his thr ser leu his lys phe phe ala
1081 CAA AGG AUC GGU UCC AAU GCU CUG GCU CGA AAC UCA CAC AUC UUU GCA CAA GUU CUU UGC

    thr asn leu gln glu asp val ile ile tyr thr leu asn asn leu glu leu thr pro gly
1141 AAC UAA CUU GCA AGA AGA UGU AAU CAU CUA CAC UUU GAA CAA CCU UGA GCU AAC UCC UGG

    lys leu asp leu gly glu arg thr leu asn tyr ser glu asp ala tyr lys arg lys tyr
1201 AAA GUU AGA UUU AGG UGA AAG AAC CUU GAA UUA CAG UGA AGA UGC CUA CAA AAG GAA AUA

    phe leu ser lys thr leu glu cys leu pro ser asn thr gln thr met ser tyr leu asp
1261 UUU CCU UUC AAA AAC ACU UGA AUG UCU UCC AUC UAA CAC ACA AAC UAU GUC UUA CUU AGA

    ser ile gln ile pro ser trp lys ile asp phe ala arg gly glu ile lys ile ser pro
1321 CAG CAU CCA AAU CCC UUC AUG GAA GAU AGA CUU UGC CAG AGG AGA AAU UAA AAU UUC UCC
```

Figure 4

```
      gln ser ile ser val ala lys ser leu leu lys leu asp leu ser gly ile lys lys lys
1381 ACA AUC UAU UUC AGU UGC AAA AUC UUU GUU AAA GCU UGA UUU AAG CGG GAU CAA AAA GAA

      glu ser lys val lys glu ala tyr ala ser gly ser lys OCH
1441 AGA AUC UAA GGU UAA GGA AGC GUA UGC UUC AGG AUC AAA AUA AUC UUG CUU UGU CCA GCU

1501 UUU UCU AAU UAU GUU AUG UUU AUU UUC UUU CUU UAC UUA UAA UUA UUU CUC UGU UUG UCA

1561 UCU CUU UCA AAU UCC UCC UGU CUA GUA GAA ACC AUA AAA ACA AAA AAU AAA AAU GAA AAU

1621 AAA AUU AAA AUA AAA UAA AAU CAA AAA AUG AAA UAA AAA CAA CAA AAA AUU AAA AAA CGA

1681 AAA ACC AAA AAG ACC CGA AAG GGA CCA AUU UGG CCA AAU UUG GGU UUU GUU UUU GUU UUU

1741 UGU UUU UUG UUU UUU AUU UUU UAU UUU AUU UUU AUU UUA UUU UAU UUU UAU UUU AUU UUU

1801 AUU UUA UUU AUU UUU UGU UUU CGU UGU UUU UGU UAU UUU AUU AUU UAU UAA GCA CAA CAC

1861 ACA GAA AGC AAA CUU UAA UUA AAC ACA CUU AUU UAA AAU UUA ACA CAC UAA GCA AGC ACA

1921 AGC AAU AAA GAU AAA GAA AGC UUU AUA UAU UUA UAG GCU UUU UUA UAA UUU AAC UUA CAG

1981 CUG CUU UCA AGC AAG UUC UGC GAG UUU UGC CUG CUU UUU AAC CCC GAA CAU UUC AUA GAA
           OPA ala leu glu ala leu lys ala gln lys lys val gly phe met glu tyr phe

2041 CUU GUU AAG AGU UUC ACU GUA AUG UUC CAU AGC AAC ACU CCC UUU AGC AUU AGG AUU GCU
     lys asn leu thr glu ser tyr his glu met ala val ser gly lys ala asn pro asn ser

2101 GGA GCU AAG UAU AGC AGC AUA CUC UUU CCC CUU CUU CAC CUG AUC UUC AUU CAU UUC AAA
     ser ser leu ile ala ala tyr glu lys gly lys lys val gln asp glu asn met glu phe

2161 UGC UUU GCU UUU CAG CAC AGU GCA AAC UUU UCC UAA GGC UUC CUU GGU GUC AUA CUU CUU
     ala lys ser lys leu val thr cys val lys gly leu ala glu lys thr asp tyr lys lys

2221 UGG GUC GAU CCC GAG GUC CUU GUA UUU UGC AUC CUG AUA UAU AGC CAA GAC AAC ACU GAU
     pro asp ile gly leu asp lys tyr lys ala asp gln tyr ile ala leu val val ser ile

2281 CAU CUC AAA GCU AUC AAC UGA AGC AAU AAG AGG UAA GCU ACC UCC CAG CAU UAU GGC AAG
     met glu phe ser asp val ser ala ile leu pro leu ser gly gly leu met ile ala leu

2341 UCU CAC AGA CUU UGC AUC AUC GAG AGG UAA UCC AUA GGC UUG AAU CAA AGG AUG GGA AGC
     arg val ser lys ala asp asp leu pro leu gly tyr ala gln ile leu pro his ser ala

2401 AAU CUU AGA UUU GAU AGU AUU GAG AUU CUC AGA AUU CCC AGU UUC UUC AAC AAG CCU GAC
     ile lys ser lys ile thr asn leu asn glu ser asn gly thr glu glu val leu arg val

2461 CCU GAU CAA GCU AUC AAG CCU CCU GAA GGU CAU GUC AGU GCC UCC AAU CCU GUC UGA AGU
     arg ile leu ser asp leu arg arg phe thr met asp thr gly gly ile arg asp ser thr

2521 UUU CUU UAU GGU AAU UUU ACC AAA AGU AAA AUC GCU UUG CUU AAU AAC CUU CAU UAU GCU
     lys lys ile thr ile lys gly phe thr phe asp ser gln lys ile val lys met ile ser

2581 CUG ACG AUU CUU UAG GAA UGU CAG ACA UGA AAU AAC GCU CAU CUU CUU GAU CUG GUC GAU
     gln arg asn lys leu phe thr leu cys ser ile val ser met lys lys ile gln asp ile

2641 GUU UUC CAG ACA AAA AGU CUU GAA GUU GAA UGC UAC CAG AUU CUG AUC UUC CUC AAA CUC
     asn glu leu cys phe thr lys phe asn phe ala val leu asn gln asp glu glu phe glu

2701 AAG GUC UUU GCC UUG UGU CAA CAA AGC AAC AAU GCU UUC CUU AGU GAG CUU AAC CUU AGA
     leu asp lys gly gln thr leu leu ala val ile ser glu lys thr leu lys val lys ser

2761 CAU GAU GAU CGU AAA AGU UGU UAU AUG CUU UGA CCG UAU GUA ACU CAA GGU GCG AAA GUG
     met

2821 CAA CUC UGU AUC CCG CAG UCG UUU CUU AGG UUC UUA AUG UGA UGA UUU GUA AGA CUG AGU

2881 GUU AAG GUA UGA ACA CAA AAU UGA CAC GAU UGC UCU
```

Figure 4

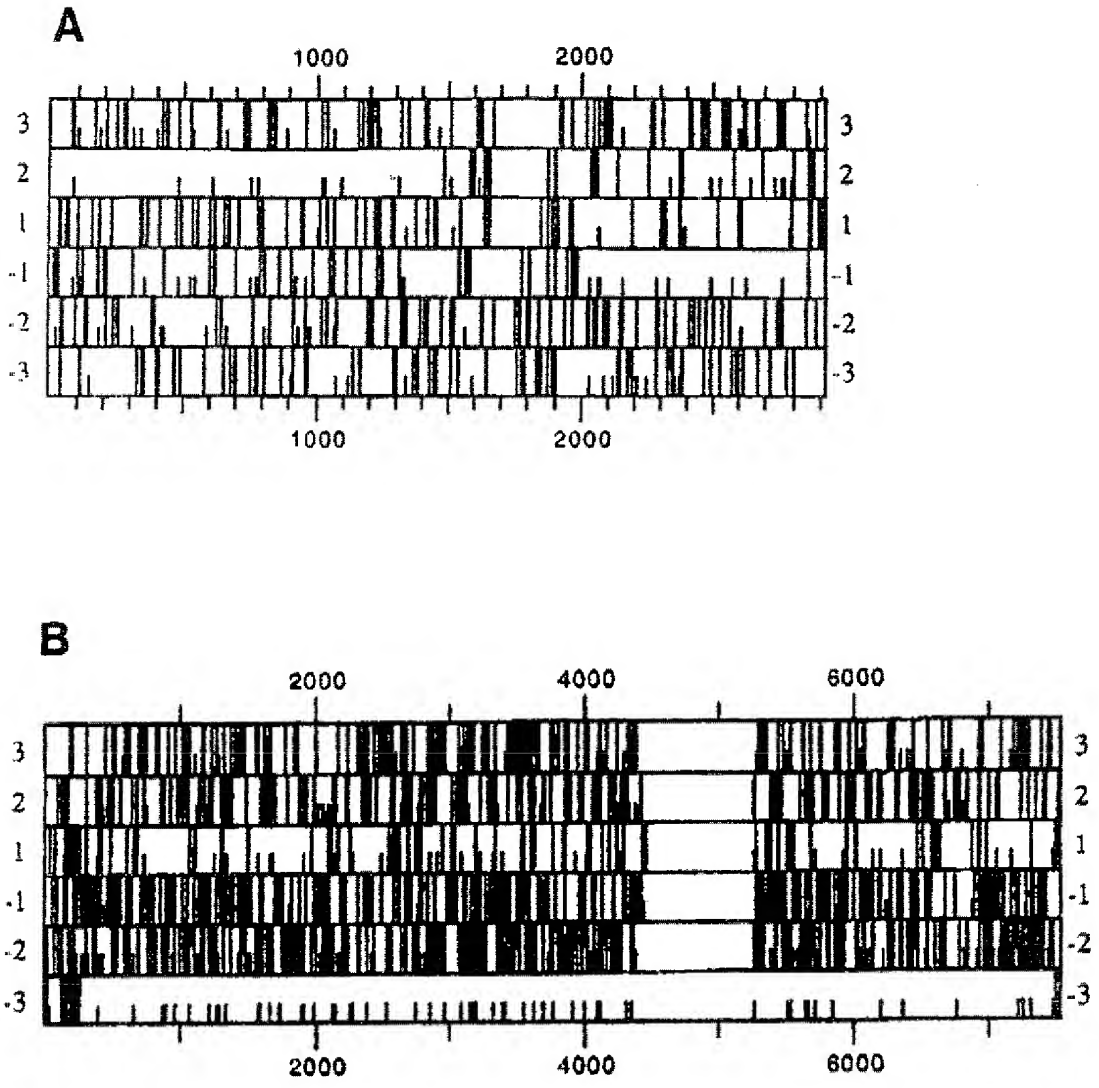
A
1000
2000
3
2
1
-1
-2
-3
1000
2000
B
2000
4000
6000
3
2
1
-1
-2
-3
2000
4000
6000

Figure 5

```
1 AGAGCAAUC. .......... ..........  .......... .......... ..........
  .......... ....CAUAGA UCAUCAAAUA  ACUUAUAUGC GAAGAUUCUG UGCAACAGUU
  UGCUAAAUGC CUGCAAUAGA GAGGAAUAAU  CGCUCCAUUU ACUAAUUCUU UCAGAUAUAA
  ACUCUACUUC UGAUGAAGAA GCAUAACUCU  UUUUUGGGUU CAAAGUUAUG CAAAAACUUU
  UGAAAUGAGC UUCAAUGCUU CUAAACAACA  UUUCUGGCAG AGAUGAGCUU GAAAAGUCUG
  UCAGCAUUUU AUCAACCUCU CCACUGGCUA  UUAAUGAUGU UGCAUUAUCA UCAGAGUGCA
  CGAUCCAUCU AGUUUGGAAA UCACAGUUUU  UGUAACAUUC CAAAGUCUUG UGAUAAGCUU
  UCAUUGCACG AGAGUGAUAA ACGGAAGACA  GAUAAUUUAA AUUGCCUUGC AACCAAUUCA
  UGCUAACAGG GUAUGUGUUU GUCGUCGAGC  CUUUGGUCGA AAGUCCUAUG GCAGUUUCAU
  UUUGCCCAAA AGUUUCUUGA GCUUUUCUUA  GAUUCAAAAA AAUAUCUGUU GGUAUACAAA
  CCUUCUUCAA UUUAACAUAC AUUAGAAUCU  UGAUGUAUUG AUUGCUCU (app. nt 5000)
```

Figure 6A

```
   1 AGAGCAAUCA GGUACAACUA AAACAUAUAA CCUCUCCACA GCCAGACUUU ACAAAUUACA  60
  61 UAAGAAUUCC CUCCAGUGAA ACUAUACCGG GACUCUACAU UUAUAUACAC UAUAGAUUUG 120
 121 GUUCACCAAA UUGUUUUAGG CUAAGUAGAU CUAAUCUAAU UCUAAAAUAG CAAUAAAUUU 180
 181 AUAUUCUAAG UUAUAUCUAA UCUGUUAUUC AUGGAUUGUU AGUCAUUACU UUCAAUUUAA 240
 241 UCUGUGUCUU CUUCUUCAUC AAGCUCAUCU UCAUCAAAGG CAUCUUCUUC CCCUCUUAGA 300
 301 UAACUAAGCA AGAAAUUGUC UUCAUCUGUU UCAACAACCU CGUCCAACUC ACUUAUUUUA 360
 361 CUCAAAAUAA UCUCUCUUGU UUGCAUGAUA UAAUCAUUGA UAUCAGAGUA GGUAUAUCUU 420
 421 CCAUACAAUG UUUUUUCUGA UAAACAUAGU GGGUGCCUAG CAAGGAAAUU UGCGAUAUCC 480
 481 UUAUAGACCU CUAGUUUUUU AGUUCGCGAU UGUUCAUCUC UUAAUUGUAU CAAAGCUUUU 540
 541 UUAACUCUUC CAGACAAGAA AUGUUCCUUG GACCUGCUUU UUGUCAGUUU UAGAUAACCU 600
 601 UCUAAGCUGU CAGCUAUCUC AUUUAAAGGA CAAUUGACAU UCUUAUAUUC CAUAUAAACA 660
 661 GUUGCAUAAA CUUUGAGUGC AGGUACACUG GAACUUAAUC UUCCCAAUCU GUCAGUUGAA 720
 721 AAUAUGUCUG GUCUAAAACU GUUCCGUAAC AAUGUCUCUA UAUCAAAGCU GGUUCCCUUU 780
 781 AUUCCUCCAA UCAAGAACUU UAUCAACUCA CUUAGUAGAA GUGUGUUUGC UGGAAAUAAA 840
 841 AACCCUUUCU UUUUCCUUUU AAUUUUCAUU UCUUUGAUCA AACCUUCAUA CAUUUCUUCC 900
 901 UCAUCUAGCC UCAACACAUU AAAACUGGCA AGACCCACAA AUUCUUCAUC AGACCUAUCC 960
 961 AUUUUUAGAU UCUCUCCAGA UAUAACAAAC CCAUUUUCUU UCAGUAUUUC CCUAGAUAGU 1020
1021 UCUGAUAAUU UUAUUUCUGG UAAAACCAGC GUCUUGAAAU CAUAACUUUU UUCCAUUAAG 1080
1081 UUGUUUUGCA AGUCCAGAUC AUGAUCCAAU CUUAUUCUCC UCCCAAAUGU GGCAUAUGUC 1140
1141 CUGCUAAGAA CAUUCCUUAA ACUUGCUAAG UUUAAUAUAG AGUCAUUGUU GUGGCAUGUU 1200
1201 UCACAAGCUU UAAUCAAUUG GAUCAUUCCU AACAAGUUGA AUUCCCCUAU CUCUAUGUGC 1260
1261 UUUUUUGCCA UUAGCUCAAC AGCAUGCAUU CCAGUGGCUC UAUAGCUUUU AAAUCCUGAU 1320
1321 GCAUCAUUUA UUAGAGAAUA UUUAAGCAUG AGCAGUAUUU CUUUGAAACC UUCUAAAUCU 1380
1381 CUAGUAUACU GUUUUGCUUU GUCAACAAUC UCUGAAAAAA GAGGACUAUA CUCAUUUUCA 1440
1441 AAGUUGUCAU GUGUUCUUAG CAAGUUUUCG AAUAGUUCUA CUGUGUCAUU AAAUCCAUUU 1500
1501 UCAUCAAGAU AGCUAUCAUA UUCUUCCACA CUUUUGCAGU UUUCCAAACC UUCUAAGCAG 1560
1561 GUAUCCCGGA UGUUCAUUAA GAAUUCAUCC AUCGUUUCUC UGAUAUCAUC AUCUACAAUC 1620
1621 UCUUUGAUCU GCUGUGAUAA GUUUCCAAAG GCUUCUAUGA AAGUUAAAGA AGUCAUAAGC 1680
1681 CUGUUUUUUA UGGUUUUGAU UGUUUCAGUU UCAGAGUCUG CAGUUUUAUA CAUAUCAGAG 1740
1741 AAAAAUUCCC CUACUUCAGG GAGCCUUAAG ACUAGAGAUU UUUGGCUUUC UAUAAUAUCC 1800
1801 CAUACAUCGG AAUUCACAUU UUCAUAUUUU AUCUUUACAA CUAAUUUUAC AUUCUCUAGU 1860
1861 CUUAUUUGUC UUGAUGUAGC AUUAACCUUU ACAAUCAUUU UGUUGUCUGU GGUCAUUCUG 1920
1921 AUGCAAGAUC UCCUGGUUCU AAGAAAACAA UCUGUGGUUG UUUGAUUUAU UAUGUUCACU 1980
1981 UUAUCAAACA UAAUUUCUGC CAAUAAUUUC UCUAUACUUU CUCUGUAUGU GUCAAAAAGA 2040
2041 UACUCUCUUU CAUUCAAUAU CUCAAAUGGU GUUCUCAAAG AAACUGCUAC AUCUUGAUCA 2100
2101 GAUGCUAUCA AUUCCACAUA GCAGCUUCUU GUCAUGAUCC CUAUUUUGUA CAUAGUUCCA 2160
2161 GAGACAUAAU UUCCUGAUUC AUCUUGUUCU UCAGUUUUUA CAUAUCUUAC AUCCCUUCUU 2220
2221 CUAAUAUACC UGAUAACAAA AUCUAAUGUU CUGUGAGAAC ACAAACCCAU GUUCCACAUC 2280
2281 CAGACAAUUU GUUUCUGGAU UGUGUUCAAU UCACUGUUAU GUGAUAGAAU CACAUUUGAC 2340
2341 ACUGUGAAGC CACUUGCUGU UGAAUUGUUA AUCUUUUUCA UUUUGUCUUC UAGACUCUGC 2400
2401 CAUAUAGGUU UACUGUCAAG GCAAACAUUU UGCAAUAUUA AAACUAUAUC CUCUCUUGUA 2460
2461 AUUGUUGGGA AAAGUGACUG CAAUUUAUCU GAUAUGUAUA CAGCAGUGUC CAGUAUUUUC 2520
2521 UCUAAGUUCA UAUCUUGCUG AGCACUUACC ACAGUUCUGA UAGUUUUGUC UAAUUUAUAA 2580
2581 CUAGCUGUAG AAACAAAAUG AUAGGAUUCU UCUUUACUGU AGAUGGUGGA UAGGAGAGCU 2640
2641 AAUGUGUUUG AGAUAAACCC UCUCCCUAGA UUGAACUUAG AAUAGAAGGG AAAGCUUUUC 2700
2701 CUAGUAAAGC AUGUGUUAAU UGCUAAAAGC CUGUUUUGUG CUGCUAUCAU AUAAUUGCAU 2760
2761 GCUUUCACCA AUCUUUGACA UUCAUUAGAA UUGCUAAAAG CCUCUGUUUC GAAGCUGUCU 2820
2821 UUGCCAUUAA CUAGGGAUUU UACAUCUUCC AUGAUUAAAC CUGUUCCAAA CGUGCUUGUU 2880
2881 AAAACAUUAA UACUAUCGUC UGUUAGAGGG AUGUUUUUUA AACCUAAGUU CUCUAUAUAA 2940
2941 UUUGAUCCAU AGACAGCAUG CAUUACUACU UAGUUAUCUC UUUUCCUAGA UUCACAAGGG 3000
3001 CAUCUCUCAG UCUGAAACCA CCUGAAAUUG UAUGCAGAAU CUAAUAUUGG AAUCAGCAUG 3060
3061 UAUGCAGAAU CUAAUAUUGG AAUCAGCAUG CUGAAAUUAA CAUCCUUGAU AACUAGAUUU 3120
3121 ACACAUGUUU UAAAUGAACA AUAUGAAUCU AAAUCAAAUU UUAGCAUUUC UAAGAUUUUC 3180
3181 UUUUCCAUGU UUUCUUCAUA AGUCAUCUUC ACAGUGGAAG AAAUGUUUAC AAAUGUUCUA 3240
3241 GCUUUUGACC UUAACACUCU AUCUAAAGCA AGUUUCUCUG UAGAUUGGUU CGAUAACUGG 3300
3301 UUUCUUUUAG AAGGGCUUGU GUACAUGAAC AGCUGACUAA GUAAAAAAUC UCUAUCAUUC 3360
3361 AAUGGUGCUA UCAGAAUUGU UUCAGGAUUU UUUAUCAUUA UGUCAUGGAC AGAUUCCAUU 3420
```

Figure 6B

```
3421 UCUAAGCUAG AUGCAAUCUU CUUAUAUAAC UCAUCUUCCU CUAAAACAGC AUCCUCGUCU 3480
3481 AUCUUUAAGU UCUUAGUGCU UUUAUAUAAA UCUUCAUUCU UUUUUAAAUU AGGAUAAAGU 3540
3541 UUUGAAAAAU CUUUAUAACU UGACAUCUUG UUUAAAGCAU UCUCAUUUAU AAAAUUAGGU 3600
3601 AAUUUUAUAA UCUGAGUUAA CAUAGUUUGA AAUUUCAUCC CAAUUGUCAG AACGUCUGAA 3660
3661 UCUUCAGACA CUGAUGCUUU CUCAAACAGA UUGAUAAGAA AUAUCAUCUU UUUAUCUUUU 3720
3721 UCUUCCAGAG UUCCUCUUUC AUACUUUUCU UUUAGCUCUC UGAAUCUCAU CUGUAGAUAA 3780
3781 CUGGAAGAAG AGACACUAUC UUUUACUUCU UCUAAACUUU UUUUGUUCAA AAAAUCUCUU 3840
3841 AUCACAUUGU AAUAGAUGAU UUGAUCAUUU GAUGAUGGAC CUAAUAUAGA CAACGGCUCU 3900
3901 AUAGGAGAGG UCAACCAGCC CCCCAUGUUU GUGGGAAUCU CGUUUGACUU UAAACUUACU 3960
3961 CCAAGCUUGU UAAAAAUUCU AAUACUAUCA UUCACUUCAC CAGGAAGCAU UGAAUAGAUG 4020
4021 CUUAACGCUU GUACCUGCAC AGCCCCAUAA GCAAAAGGUA UAACUUCAUU AGGACAGCCU 4080
4081 UUUCUCAGAA GCAUUGUAAC AUGUAUACUG AGUGACAUUA GAUCAUCAAA AUAACUUAUA 4140
4141 UGCGAAGAUU CUGUGCAACA GUUUGCUAAA UGCCUGCAAU AGAGAGGAAU AAUCGCUCCA 4200
4201 UUUACAAUUC UUUCAGAUAU AAACUCUACU UCUGAUGAAG AAGCAUAACU CUUUUUUGGG 4260
4261 UUCAAAGUUA UGCAAAAACU UUUGAAAUGA GCUUCAAUGC UUCUAAACAA CAUUUCUGGC 4320
4321 AGAGAUGAGC UUGAAAAGUC UGUCAGCAUU UUAUCAACCU CUCCACUGGC UAUUAAUGAU 4380
4381 GUUGCAUUAU CAUCAGAGUG CACAAUCCAU CUAGUUUGGA AAUCACAGUU UUUGUAACAU 4440
4441 UCCAAAGUGU UGUGAUAAGC UU..... ..... ..... ..... .....
      ..... ..... ..... ..... ..... .....
      ..... ..... ..... ..... ..... ...GAAUUGC
      AUCAUGUCUU CAGAGUAAUG GUAAAUGGCU AGAUCAUUGC AUAAUUUGUU UAUCUUUUUG
      CAAGUCUCUA AUUCAGUCUU CAGUACUGAA UCUUGAAUGU ACAAGCAUUC AGUAUUUUUA
      GCAUAAUUAU UGAAAACAGA UGUGAUCUUA CCAUCCUCUA UUAUAUUAAA AUUAUGAAGU
      AGAAAAAAAG AGCCCUUAAA UCUUUUACCA CUUAAGCUAU CAUCAUUGUC AUGCCUUAUC
      AGUUUUUUCA UAUGAGUCAG AGAAGACAUG UCUAUACUUG AGUAAUCAUC AUUUGGGUCA
      GCCUGCCUUU CAGAGGUAUC AUAUUCCAAG CCUUUAGAGA UCCUUUCUUU UAUUUCUUCU
      UUUUUUAGAA UUCUUUUCUU AGAACUCAUU CCUAUAUUCA UUACUGCUAA GCCUUUAAAA
      UAGGUAGUUU UGUUCUCCCU GCCUGUAUCU UUCAUGCAAA AUUCAUUUUU GUUUCUGGUU
      AAACACGAAC CUUGCUUAUU CCAUUCUUUA GACUUUUUUG UUCUCUCAAC UUUGAUACUC
      CCGGUGUUUA UAUCUAUUUC AUCUUCAAGU UCCUGCUUUU CCAUUUUCGU UGUAUCACAA
      CUUAGCUCUG CUAAACACAG UCUAACUAAG UUUGUUAUAU CUCUUGCUUU AUUUGUUUGG
      UUAUUUAUGC UACUAUAUGG GCUAUUAUAA AUAACCCUUG AAGAACUUUU GAAGCUCUCC
      CGAAACUUGU UAAUAAUUGC GUUGUUGUAA UCUACCUCUG AGCUCAAUUC UAUCUUCAAA
      CUGGGGACCA AGUCAUGGAA AGUUAGUUUU AGUUCAGAGA GAUCAUUGAA ACAGUUAGGA
      UACCUUUCUU CAUAUUCAGC AGCAUUCUGG UUAAAAGCUG UGGUUAUGUC AGGUAUCUCU
      UUUUCUUUUC UCUUUUUUUC UCUAAUUCUA AGCAUGAUCU UCUCAAGGUG UUUUCCAUAC
      AGAGAUUCCC AUAUUUGCUG CUUUUGGAUU AAGUGGUCUU UCAAAUCAGC UAAAUCAGCU
      UUUGGCUCUA UUUUUAAAUU GAUACAAGAG UUUUGGUACC CUUUAAAUUU GUCAGCUAUA
      UGCUGGUCAU UCAUUGUUUC AAUGUGGAAA CAGAAAUUUC CAAACAAAAA AUGUAUUAUU
      UCCUCUCUAG GGAUUUCAUC UGUAACAAGU GAAUCAAUUA UUCUUUCAAU AUCUUCGGAA
      UAGGCUAUCU CAGUAGGAUU UCUUUCUUUG UAUUUGCUGG UCCAGUGACU GAAAAACACA
      UCUUUAGAUC CAAAAUCAUC ACCUAAGAUU CCUCUGACUU CAGAUUUGAU UCUUUGAGUG
      UUUGCAUAAU AGUUUCUCAC CACUGUGGGU GUGGUUCCAA CUAUUUGCCC AACAUAAGAA
      AUAUUUAGCA AUUUGGCAUU UUUAGCAUGC UUUAUUCUUU CUCCGUGCAA UGCUGAUGGA
      GAAUCUUCGA AUUUCUCCAA AACCCUGCUA CUAGUAACAG UGAUUGGCAU CCCAGCAAUA
      GGUUUCAAUA UAACUAUAGA AACGAAAACA GGAUGCUCUU CCAGGAAUGG CUUUCCAUUC
      ACUUUUAUAU CUUUGAAAGA UUUCCAAAUA UUCUUGUAGG UAUUUCUCCA UUGUUUGGUU
      UCAGUCCUGG CAUCGACAGA UAUUUUCCAA UCAUAAAUGA UUAAACAAAG CUCACUGUUU
      UUUGAUUCUU UAUAUAUCAC AUAGUUAUCA GGGGUGGCAU UAGUUGCUAC AUAUUCUUUA
      UUUAUUCUCU CAAACUCAUC AUCAGAGAGC UUGUUAAUGC AGGACUUUUU GUUGAGUUCA
      UCCAGAUAUU CUCUUAGUGC AUGCUCUAUC UCCACUUCAU UUCUUUGUUU GGGCUUUAUU
      CUCAGGUGCC UGGAGAUCAG CUCUCCAAAA AUGUCAUGCC UUGCUAACUC UGUUUCUAGG
      UAUUUUUGCU CAAACAGAGA GACCAUCUCA CUUAGCUUUU UGACAUCCAC AGUUGCCAUC
      CCUUUGUUUA GUCCUUCACC AUCAGCAGUG AUUUUGACAA UUAACUCUCU CAUUGUCUCA
      UAAUUUUUUG CAUUAUUAUU UAUAAUCACA UCUUCUGGGA UCUCAUCACU AUUUCUCUUA
      UGUAAAUCCA AAGCUAGAUC GUGGUUAGAA CCUACACAAU CCUCAAUAGA CAACAGUAAA
      GUGGUUCCAU UUUCUAUUAA UUUUUGUAUU UUCUGGAUGU UCAUGUUUGC UUAAAAUCGU
      UGUUACCUGA UUGCUCU (app. nt 7500)
```

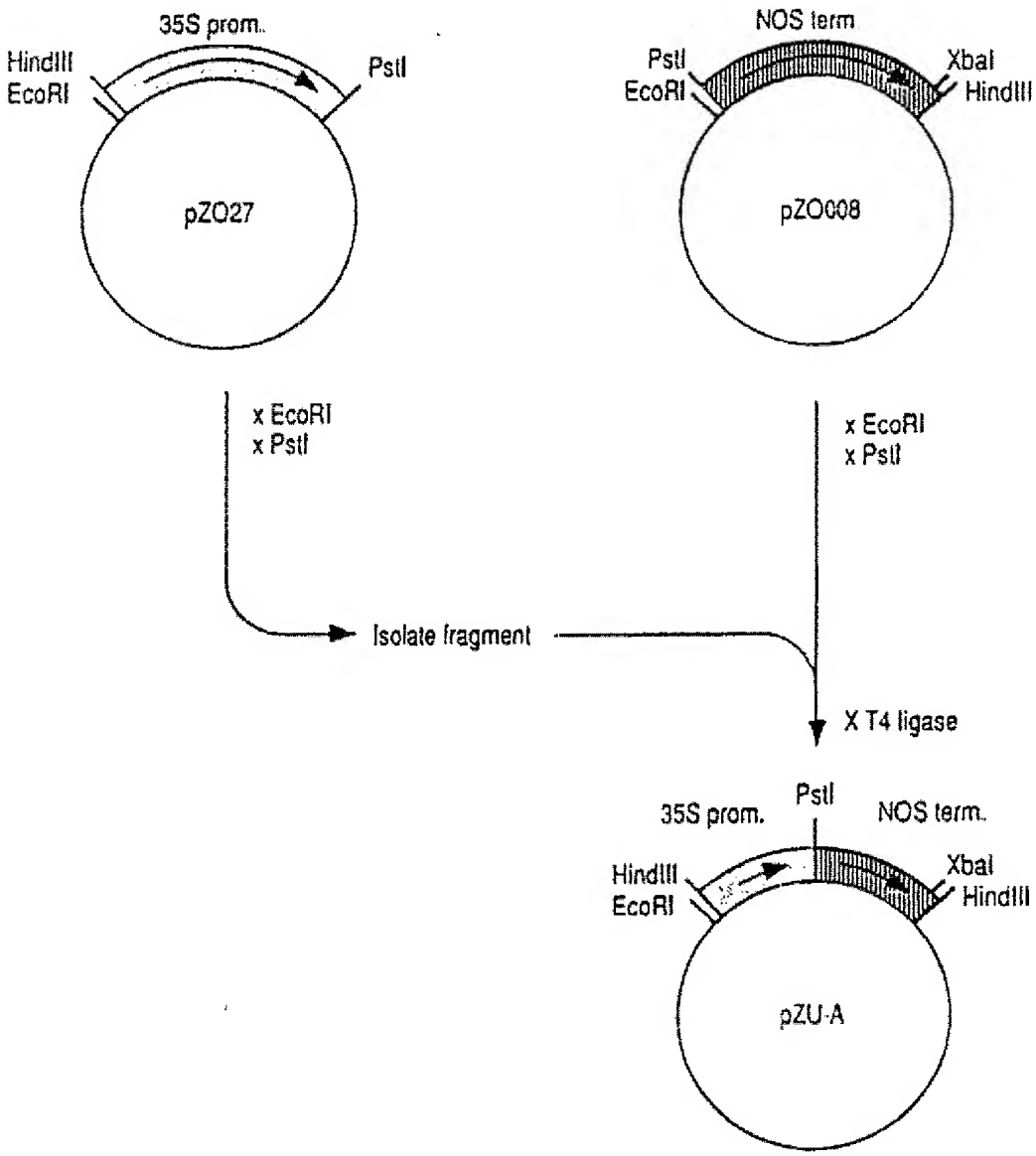

Figure 7 Construction of expression vector pZU-A

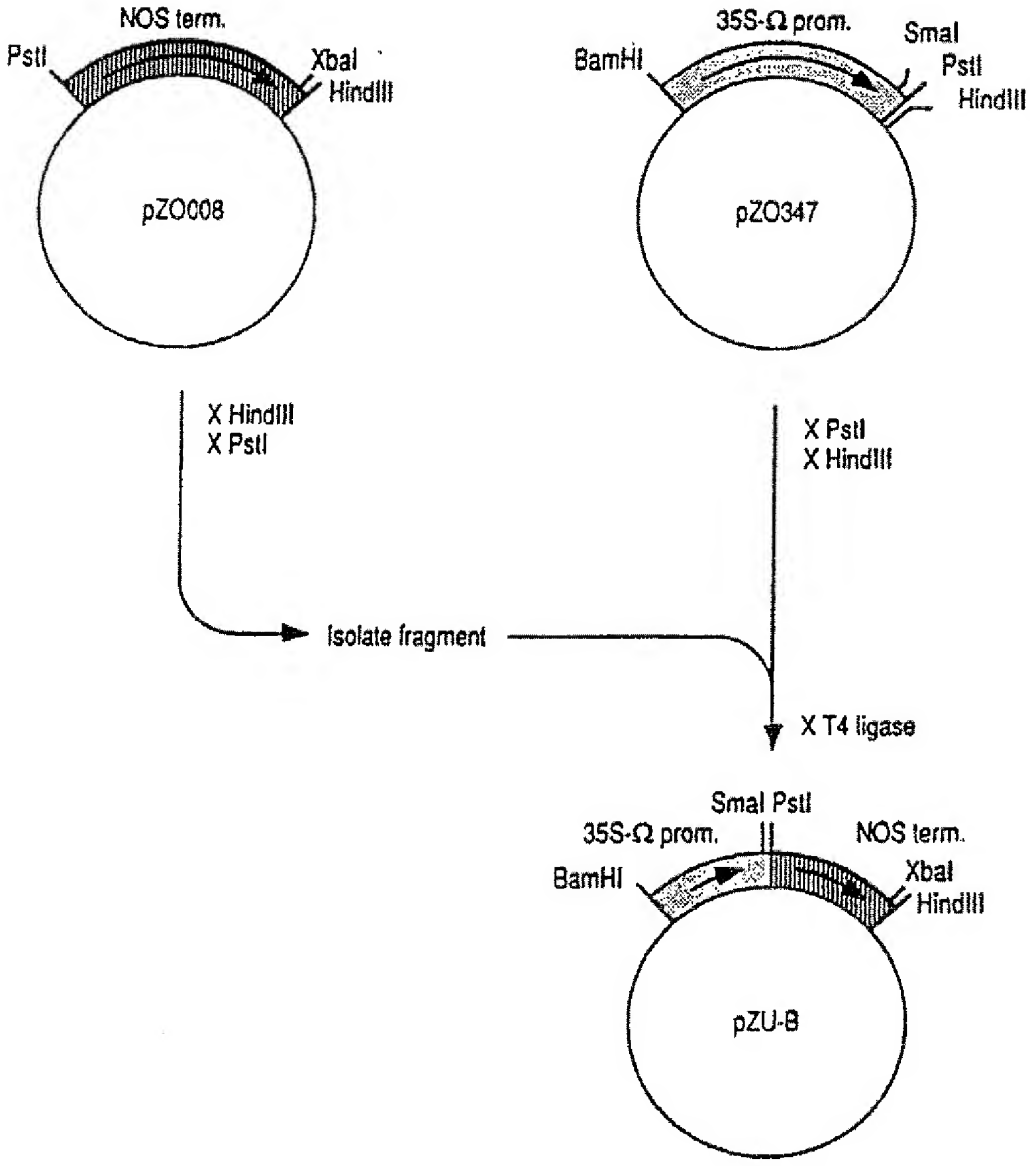

Figure 8 Construction of expression vector pZU-B

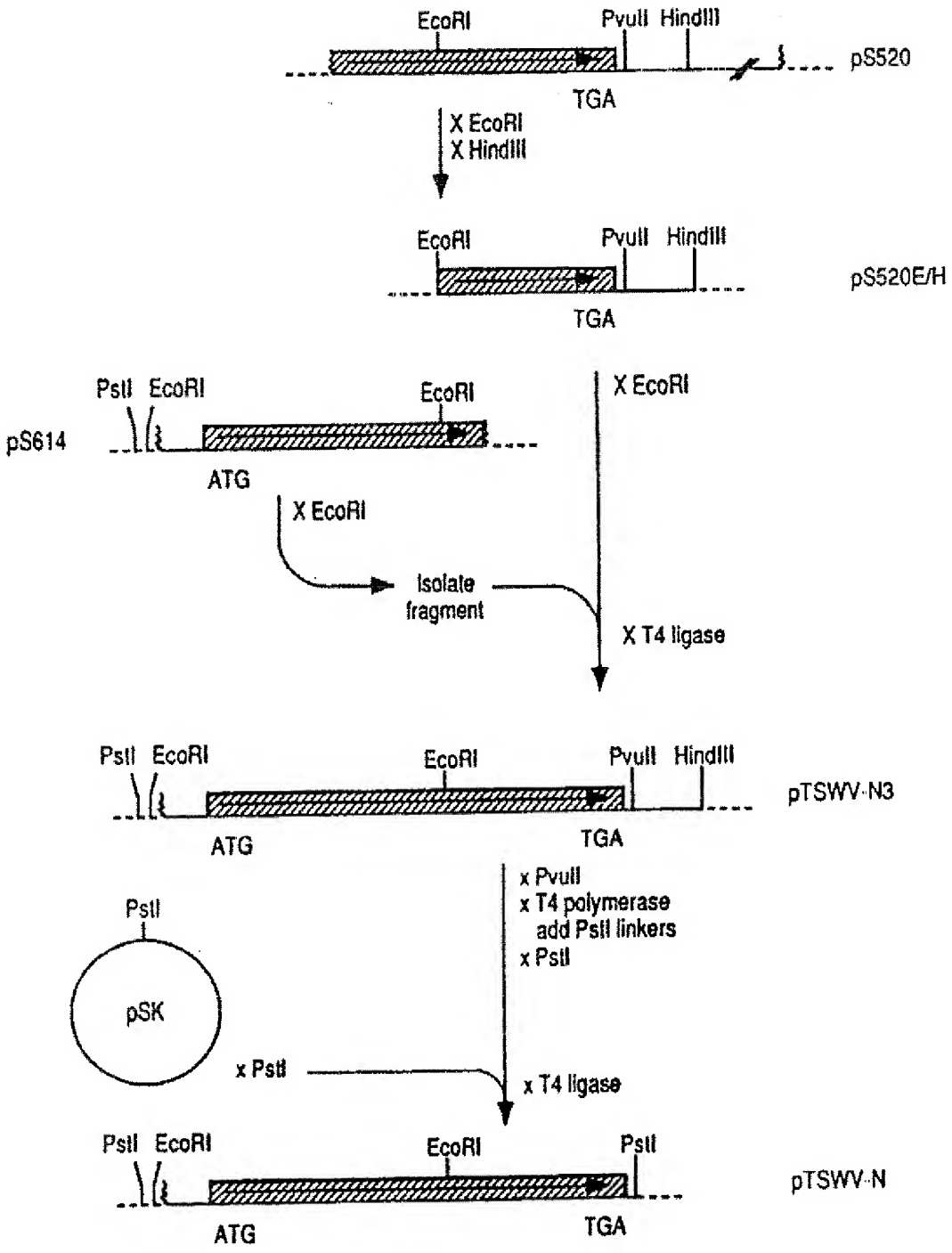

Figure 9 Construction of TSWV-N protein gene

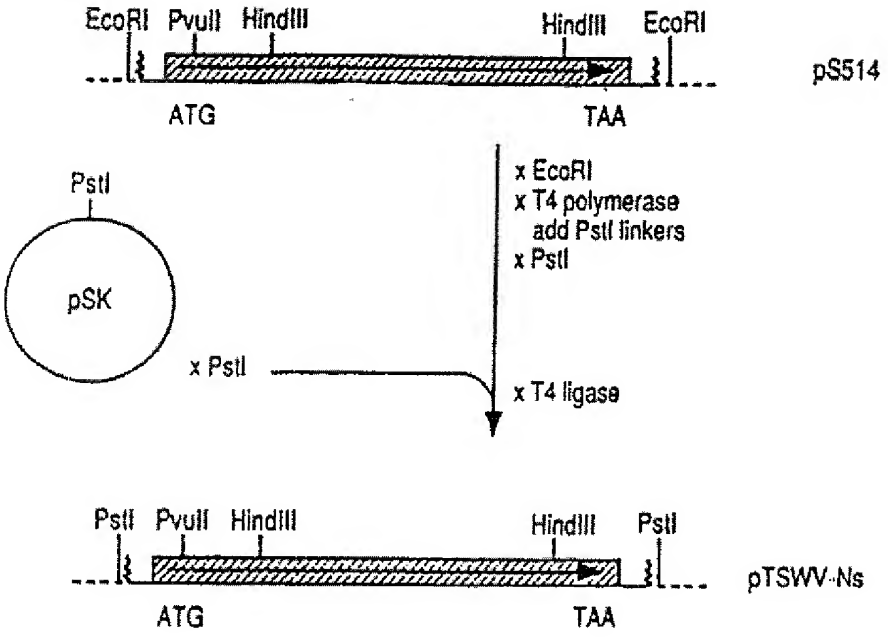

Figure 10 Construction of TSWV-Ns protein gene

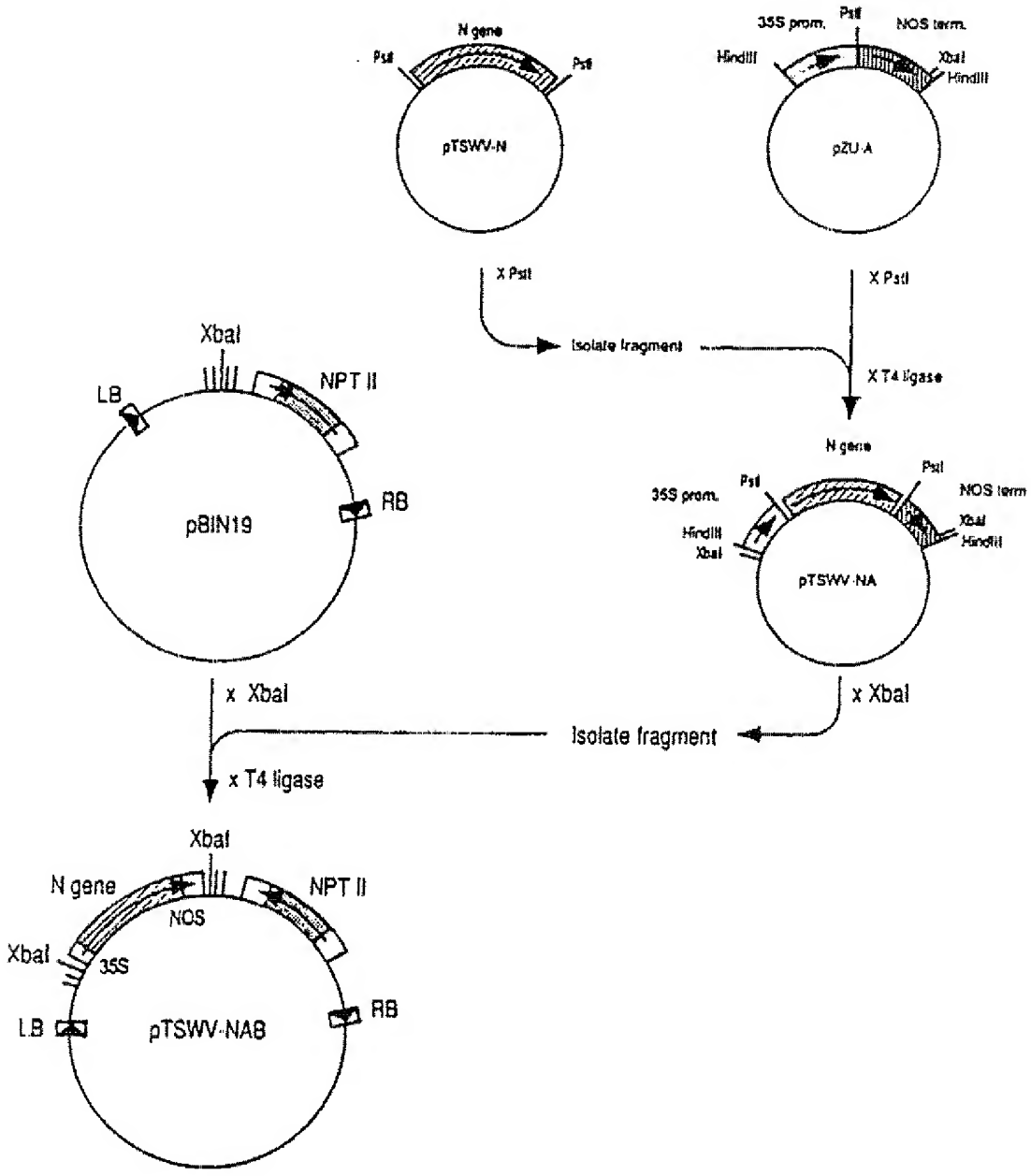

Figure 11 Construction of plant transformation vector pTSWV-NAB

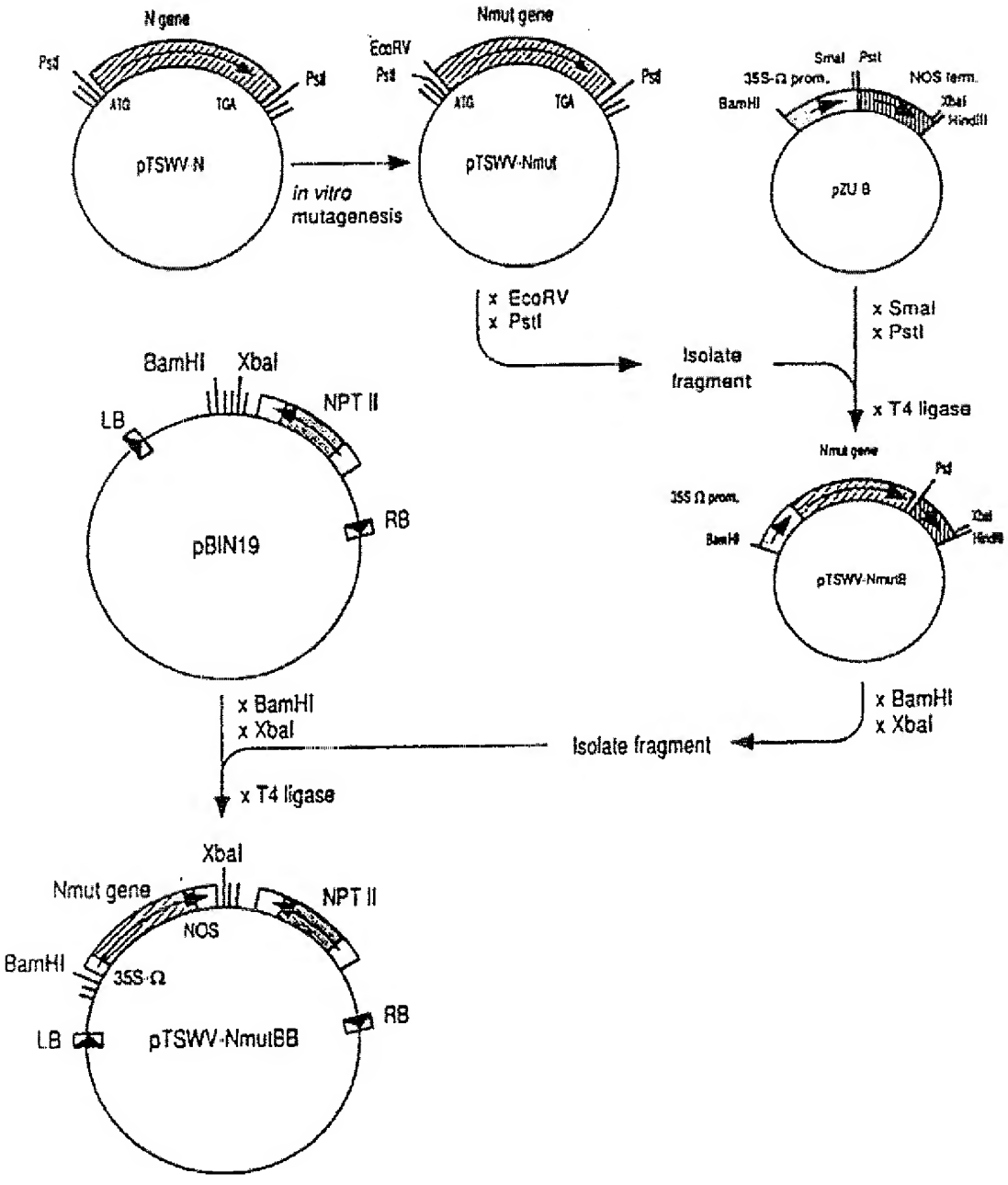

Figure 12 Construction of plant transformation vector pTSWV-NmutBB

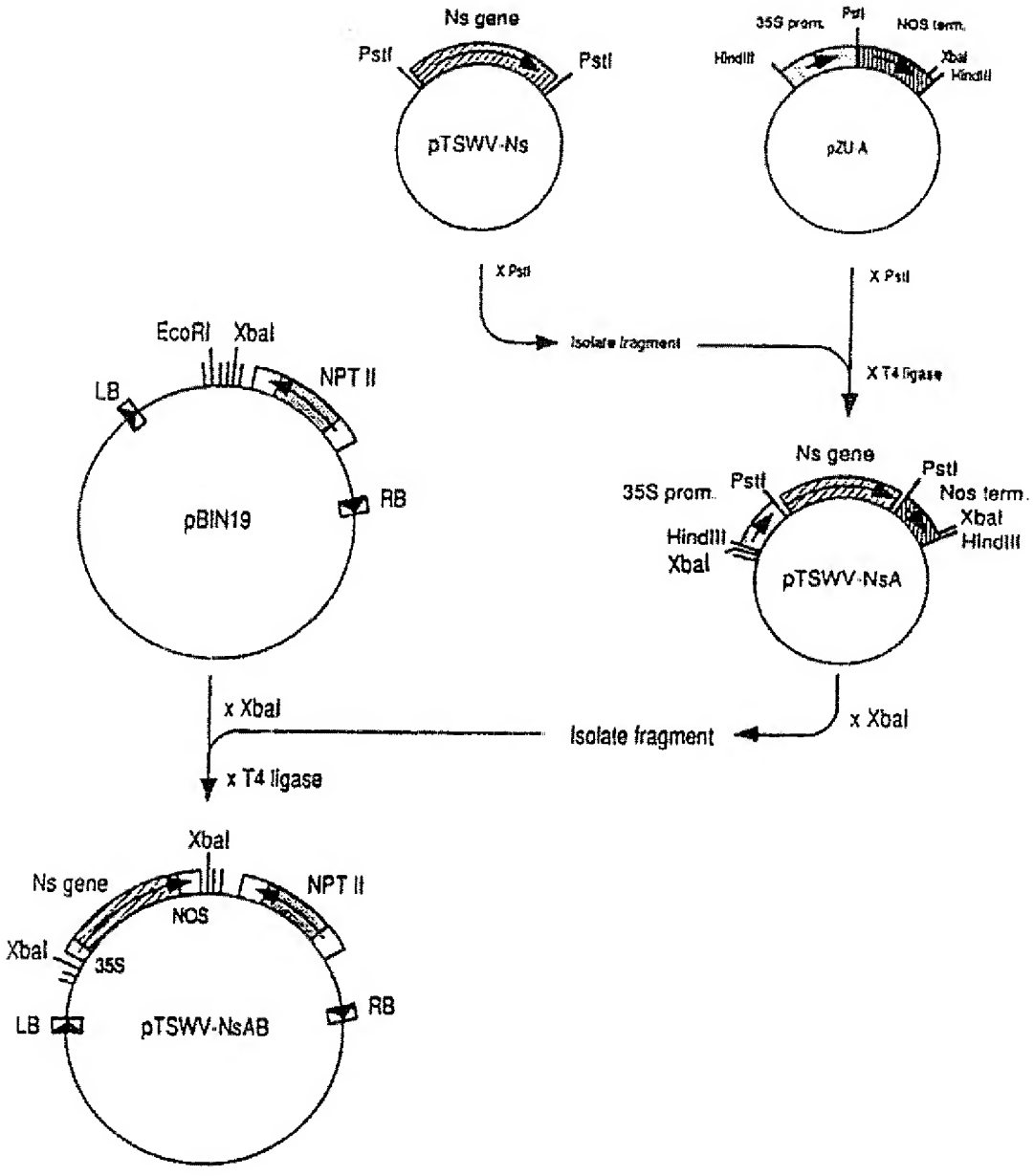

Figure 13 Construction of plant transformation vector pTSWV-NsAB

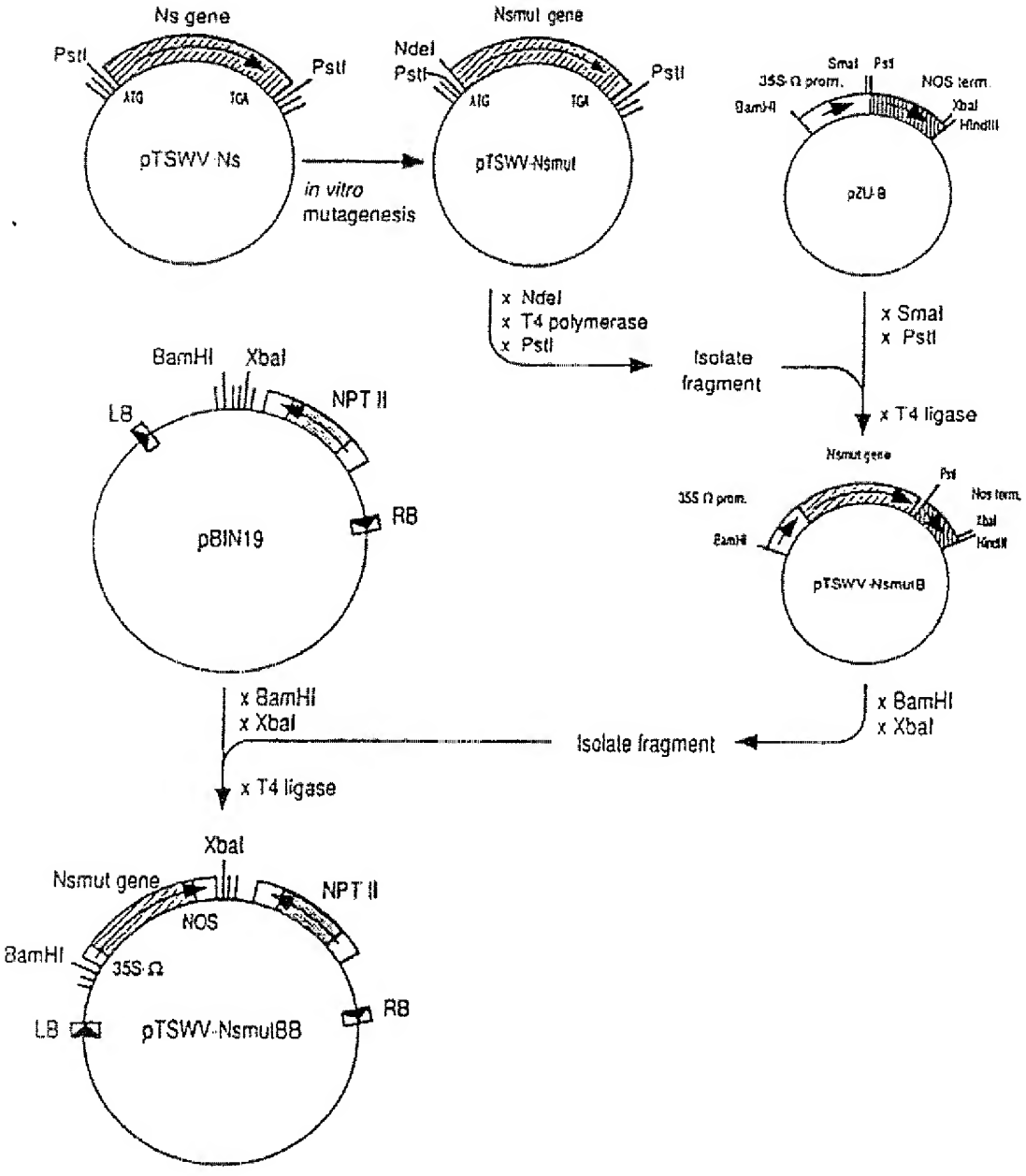
Ns gene
PstI
PstI
ATG
TGA
pTSWV-Ns
in vitro mutagenesis
Nsmut gene
NdeI
PstI
PstI
ATG
TGA
pTSWV-Nsmut
x NdeI
x T4 polymerase
x PstI
SmaI PstI
35S-Ω prom.
NOS term.
BamHI
XbaI
HindIII
pZU-8
Isolate fragment
x SmaI
x PstI
x T4 ligase
BamHI XbaI
NPT II
LB
RB
pBIN19
Nsmut gene
PstI
35S Ω prom.
Nos term.
BamHI
XbaI
HindIII
pTSWV-NsmutB
x BamHI
x XbaI
Isolate fragment
x BamHI
x XbaI
x T4 ligase
XbaI
Nsmut gene
NPT II
NOS
BamHI
35S-Ω
LB
RB
pTSWV-Nsmut88

Figure 14

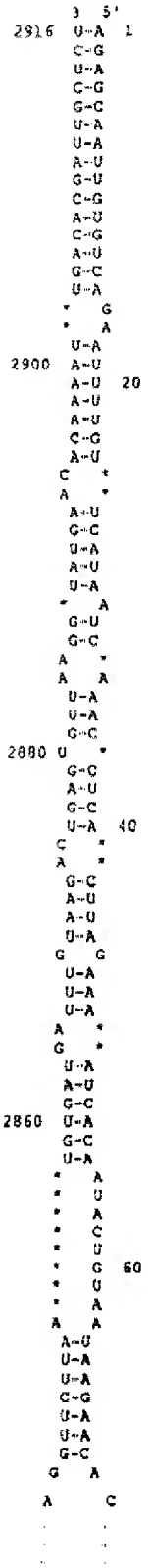

Figure 15

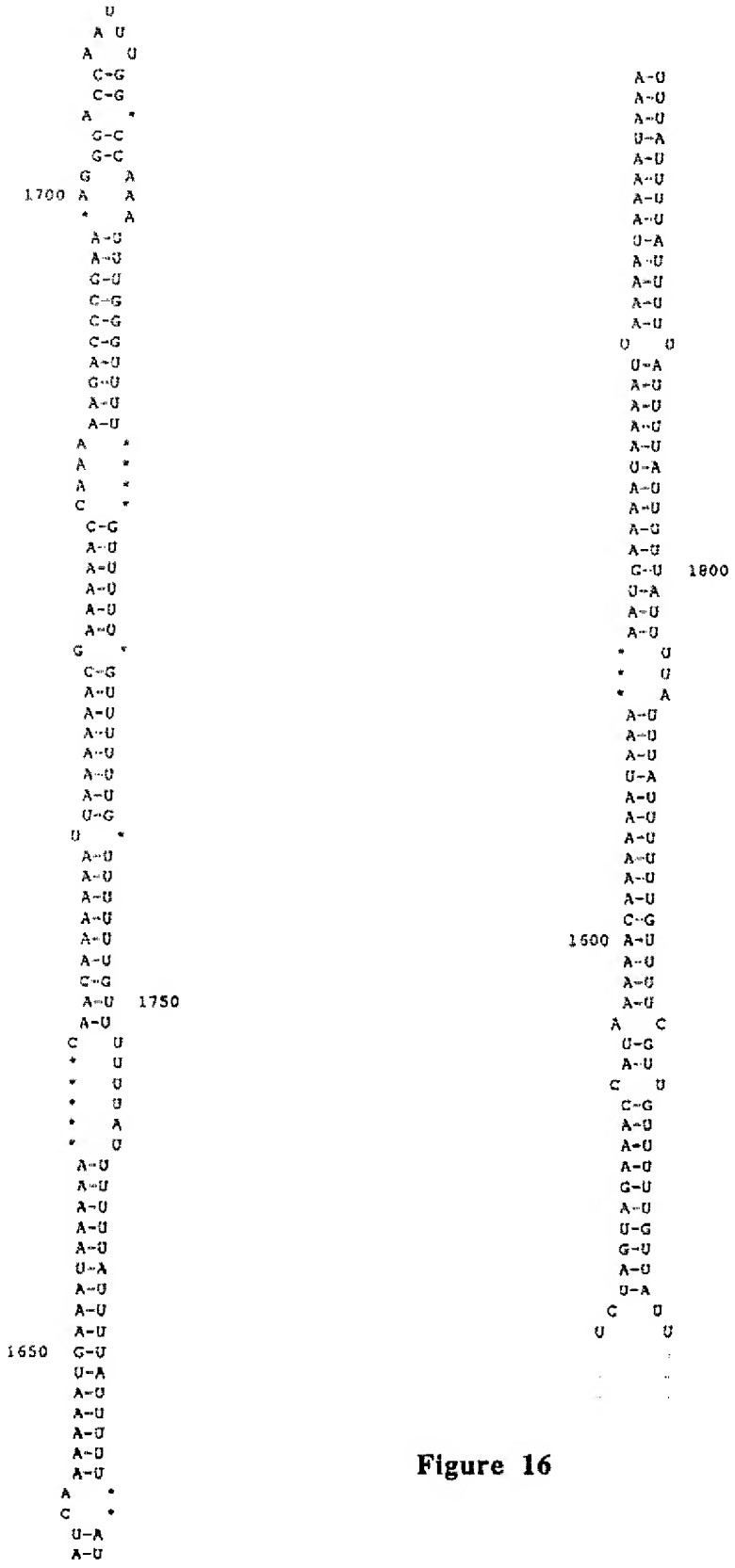

Figure 16

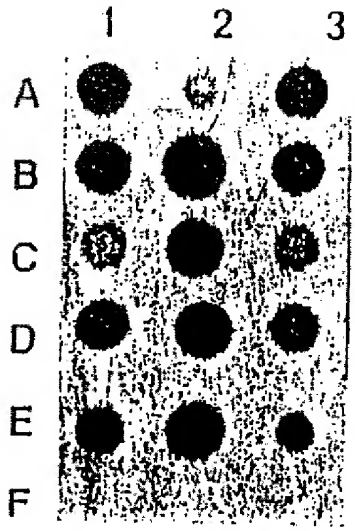

Figure 17

European Patent Office

EUROPEAN SEARCH REPORT

Application Number

EP 90 12 1052

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int. Cl.5)
O,X	BIOLOGICAL ABSTRACTS, BR36:106479, Philadelphia, PA US; T L. GERMAN et al : "Dot blot detection of tomato spotted wilt virus using cloned complementary DNA probes", & PHYTOPATHOLOGY 78, (12 part 1), 1988 1599 * Abstract *	12-14	C 12 N 15/40 C 12 N 15/82 C 12 N 15/11 A 01 H 5/00 C 12 Q 1/68
X	BIOLOGICAL ABSTRACTS, vol 89, no 5, 1990, abstract no 52485, Philadelphia, PA, US; A E E. RONCO et al : "Cloned complementary DNA probes for the detection of tomato spotted wilt virus", & PHYTOPATHOLOGY 79(11): 1309-1313, 1989 * Abstract *	12-14	
D,A	EP-A-0 223 452 (MONSANTO) * Whole document *	1,15	
A	EP-A-0 240 332 (LUBRIZOL) * Whole document *	1,15	
P,A	J GEN VIROL , vol 70, no. 12 December 1989, pages 3469-3473, SGM, GB; P. DE HAAN et al : "Molecular cloning and terminal sequence determination of the S and M RNAs of tomato spotted wilt virus" * Whole document *	1,2	TECHNICAL FIELDS SEARCHED (Int. Cl.5) C 12 N A 01 H C 12 Q
P,A	J GEN VIROL , vol 71, no. 5, May 1990, pages 1001-1007, SGM, GB; P DE HAAN et al : "The S RNA segment of tomato spotted wilt virus has an ambisense character" * Whole document *	1,2	
P,X	BIOLOGICAL ABSTRACTS, vol 89, no 11, 1990, abstract no 121862, Philadelphia, PA, US; C HUGUENOT et al : "Detection of tomato spotted wilt virus using monoclonal antibodies and riboprobes", & ARCH VIROL 110(1/2): 47-62 * Abstract *	12-14	

The present search report has been drawn up for all claims

Place of search	Date of completion of search	Examiner
The Hague	07 February 91	MADDOX A.D.

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same catagory
A : technological background
O : non-written disclosure
P : intermediate document
T : theory or principle underlying the invention

E : earlier patent document, but published on or after the filing date
D : document cited in the application
L : document cited for other reasons

&: member of the same patent family corresponding document